

Supplementary Information

Continuous evolution of SpCas9 variants compatible with non-G PAMs

Shannon M. Miller^{1,2,3,6}, Tina Wang^{1,2,3,6}, Peyton B. Randolph^{1,2,3}, Mandana Arbab^{1,2,3}, Max W. Shen^{1,2,3,4}, Tony P. Huang^{1,2,3}, Zaneta Matuszek^{1,2,3,5}, Gregory A. Newby^{1,2,3}, Holly A. Rees^{1,2,3}, and David R. Liu^{1,2,3,*}.

¹Merkin Institute of Transformative Technologies in Healthcare, The Broad Institute of Harvard and MIT, Cambridge, MA, USA

²Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA

³Howard Hughes Medical Institute, Harvard University, Cambridge, MA, USA

⁴Computational and Systems Biology Program, Massachusetts Institute of Technology, Cambridge, MA, USA

⁵Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA

⁶These authors contributed equally

* Corresponding author

Supplementary Figures.

1. Additional details of Cas9 DNA-binding PACE and Cas9 nuclease selections.
2. Effects of mutations on PAM recognition by SpCas9 variants.
3. Characterization of Cas9 variants evolved in this study, and SpCas9-NG in bacterial PAM depletion experiments
4. Characterization of PAM preferences using a genomically integrated 11,776-member human cell target sequence library.
5. Characterization of evolved variants, SpCas9-NG, and SpCas9 in human cell indel formation experiments.
6. Characterization of off-target activity of evolved Cas9 variants.
7. Characterization of evolved variants, SpCas9-NG, and SpCas9 in human base editing experiments.
8. Sickle-cell anemia allele editing by ABEs derived from evolved SpCas9 variants.

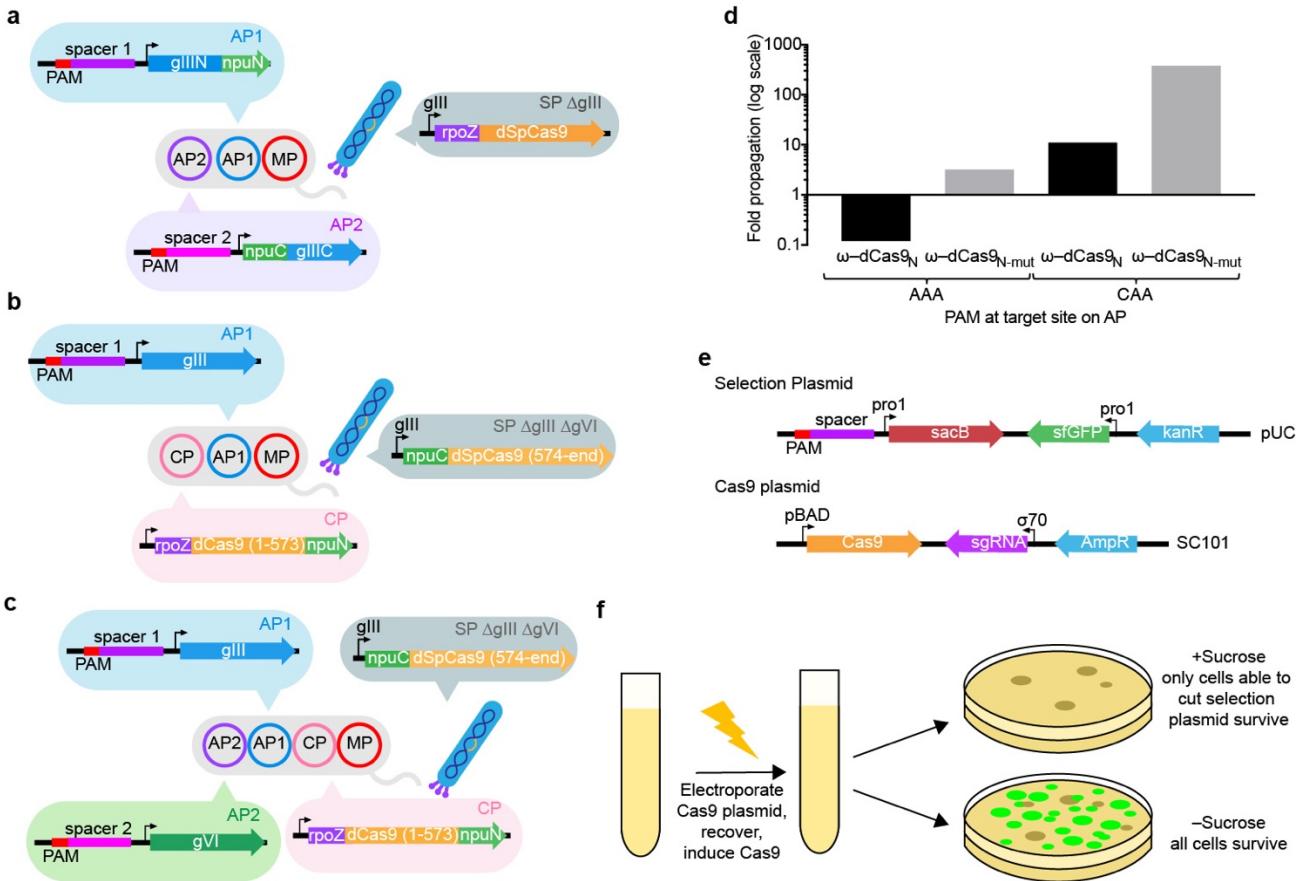
Supplementary Tables

1. PANCE-evolved variant mutation table
2. PACE-evolved variant mutation table
3. Bacterial PAM depletion scores
4. GUIDE-seq identified off-target sites
5. Genomically integrated 11,776-member human cell target sequence library editing data
6. Sample sizes (n) in Figure 3b.
7. Plasmids and selection phage (SP) used in this work
8. Primers used in this work
9. Protospacer sequences used in this work

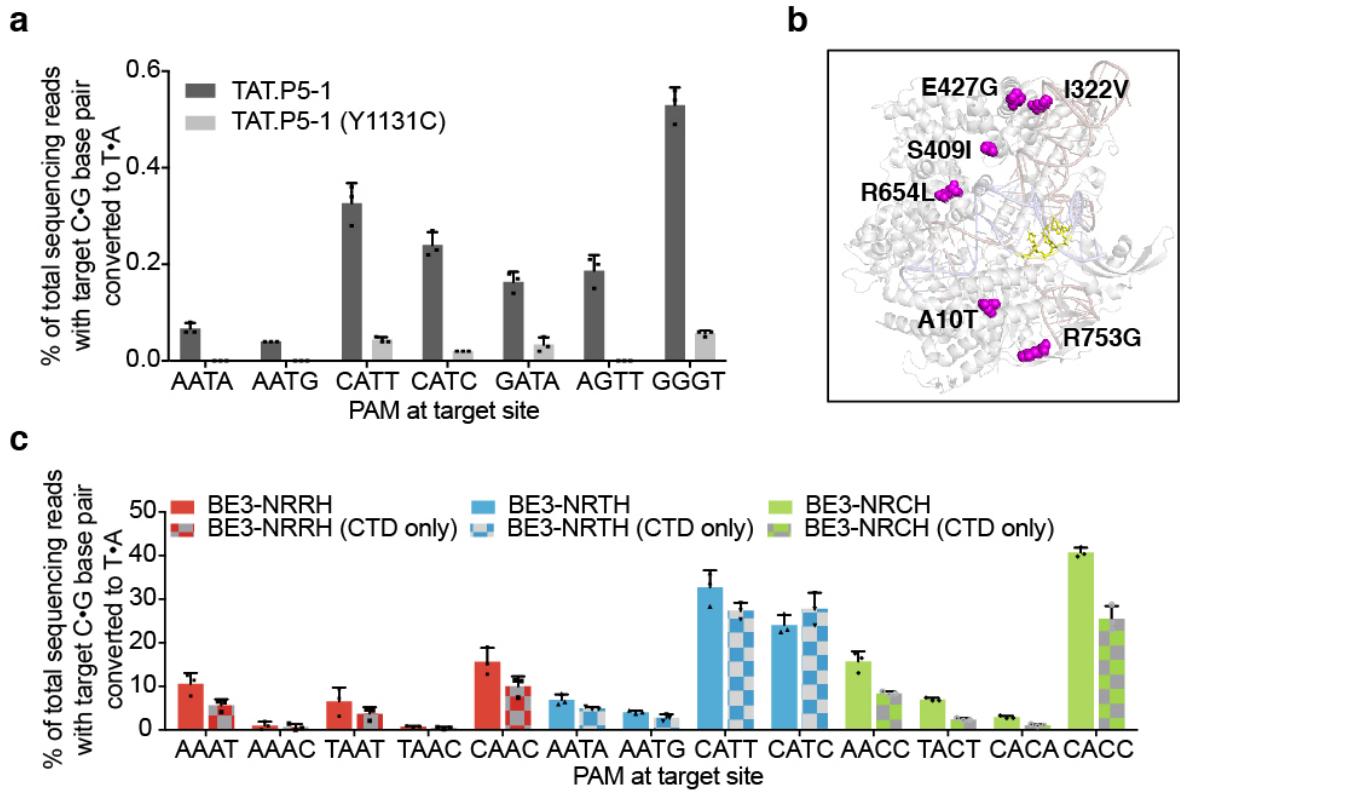
Supplementary Note

1. Evolution of SpCas9 variants that recognize NAT or NAC PAM sequences
2. Bacterial PAM depletion
3. sgRNA preferences of evolved variants and SpCas9-NG in mammalian cells
4. Evolved SpCas9 amino acid sequences

Supplementary References

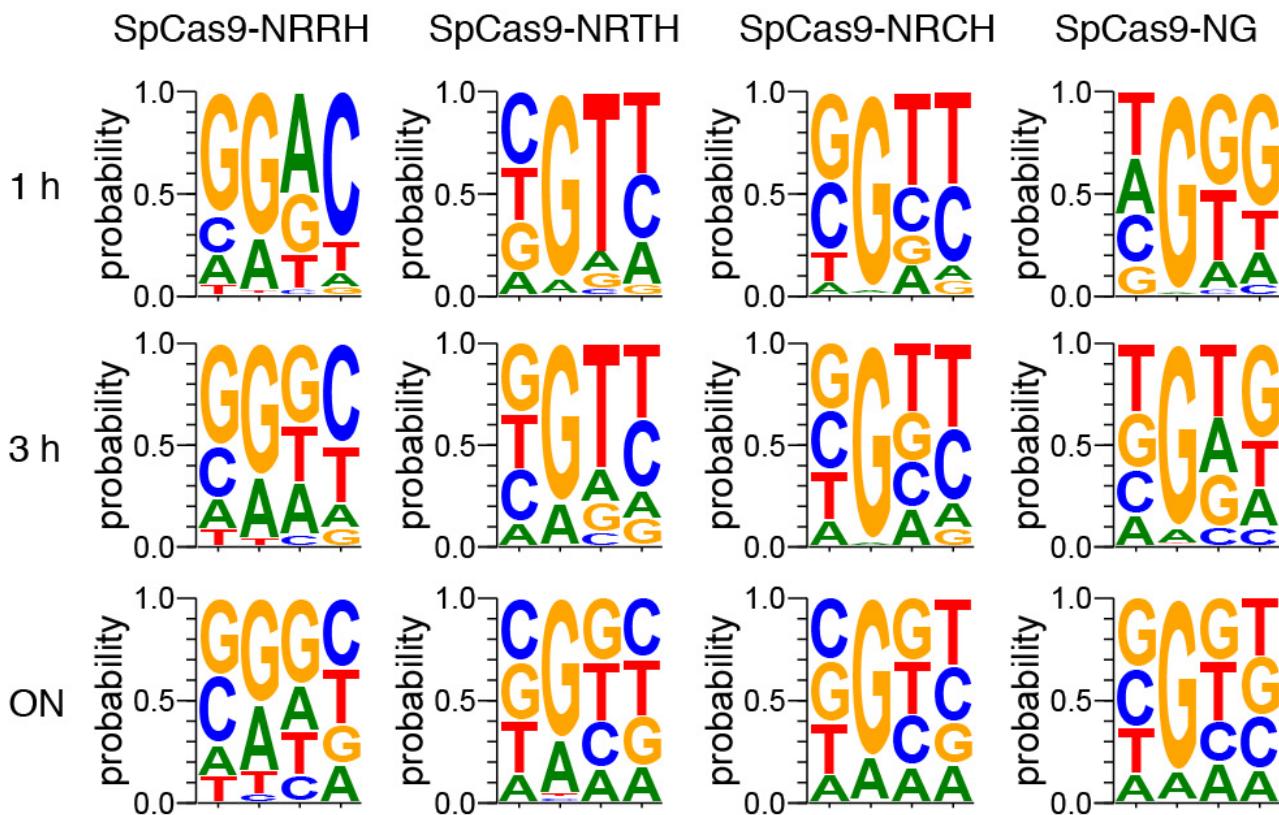


Supplementary Figure 1. Additional details of Cas9 DNA-binding PACE and Cas9 nuclease selections. (a) Dual AP selection in which ω -dSpCas9 binds two distinct protospacer-PAM sequences to drive expression of each half of split-intein pIII. (b) Split-intein Cas9 limits total Cas9 concentration in host cells, thus avoiding saturation of protospacer-PAM binding sites. Residues 574-1368 of Cas9 fused to NpuC is expressed by Δ gIII SP and ω -dSpCas9(1-573) fused to NpuN is encoded on a low copy complementary plasmid (CP) in host cells. (c) Combination of the selection principles from (A) and (B) through use of gVI as an additional PACE-compatible selection marker for phage propagation and Δ gIII Δ gVI SP. (d) Overnight propagation assay of selection phage (SP) encoding dSpCas9_C on host cells containing a complimentary plasmid (CP) providing either ω -dSpCas9_N or ω -dSpCas9_{N-mut} and an AP encoding either a AAA or CAA PAM. (e-f) Scheme of survival-based selection for Cas9 nuclease activity. Cells containing a high-copy selection plasmid encoding a protospacer-PAM sequence, sfGFP, and the conditionally lethal protein SacB are transformed with a library of nuclease-active Cas9s encoded on a low-copy plasmid that also includes the matching sgRNA. Binding and cleavage of the designated protospacer-PAM by Cas9 leads to destruction of the selection plasmid, resulting in loss of both sfGFP and SacB expression, allowing cells to survive on sucrose-containing media.

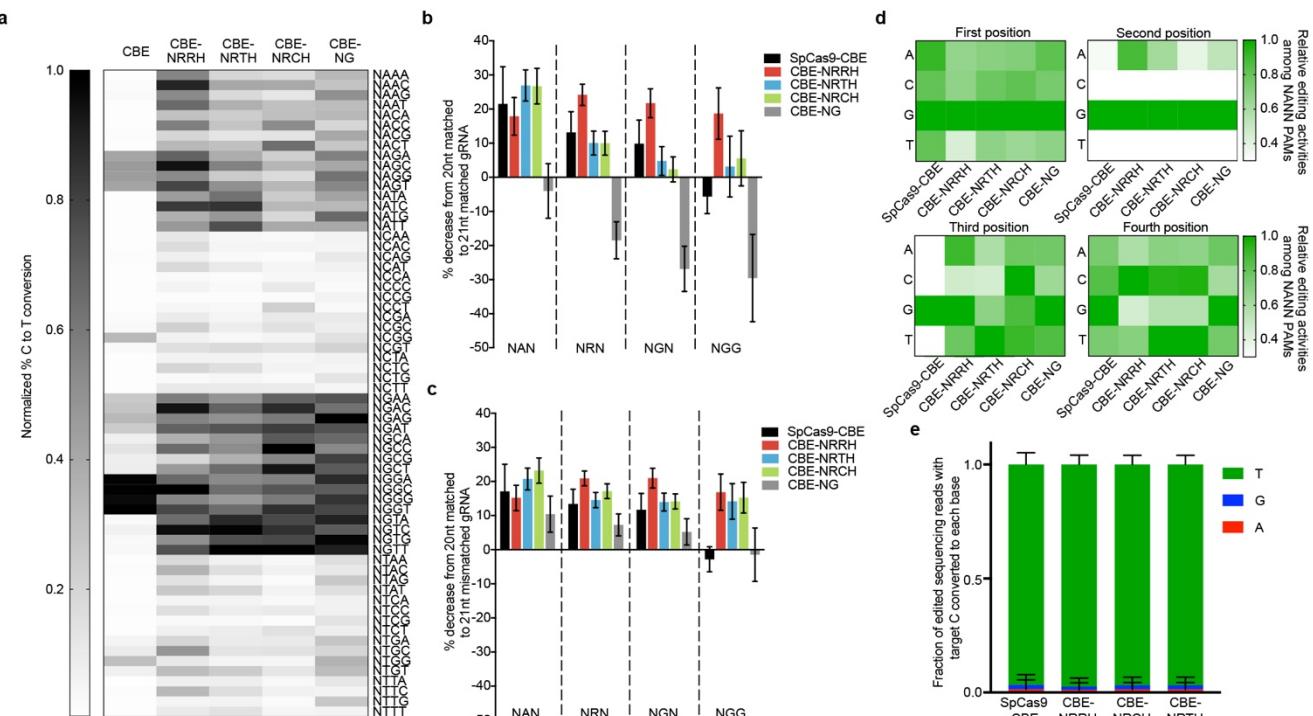


Supplementary Figure 2. Effects of mutations on PAM recognition by SpCas9 variants.

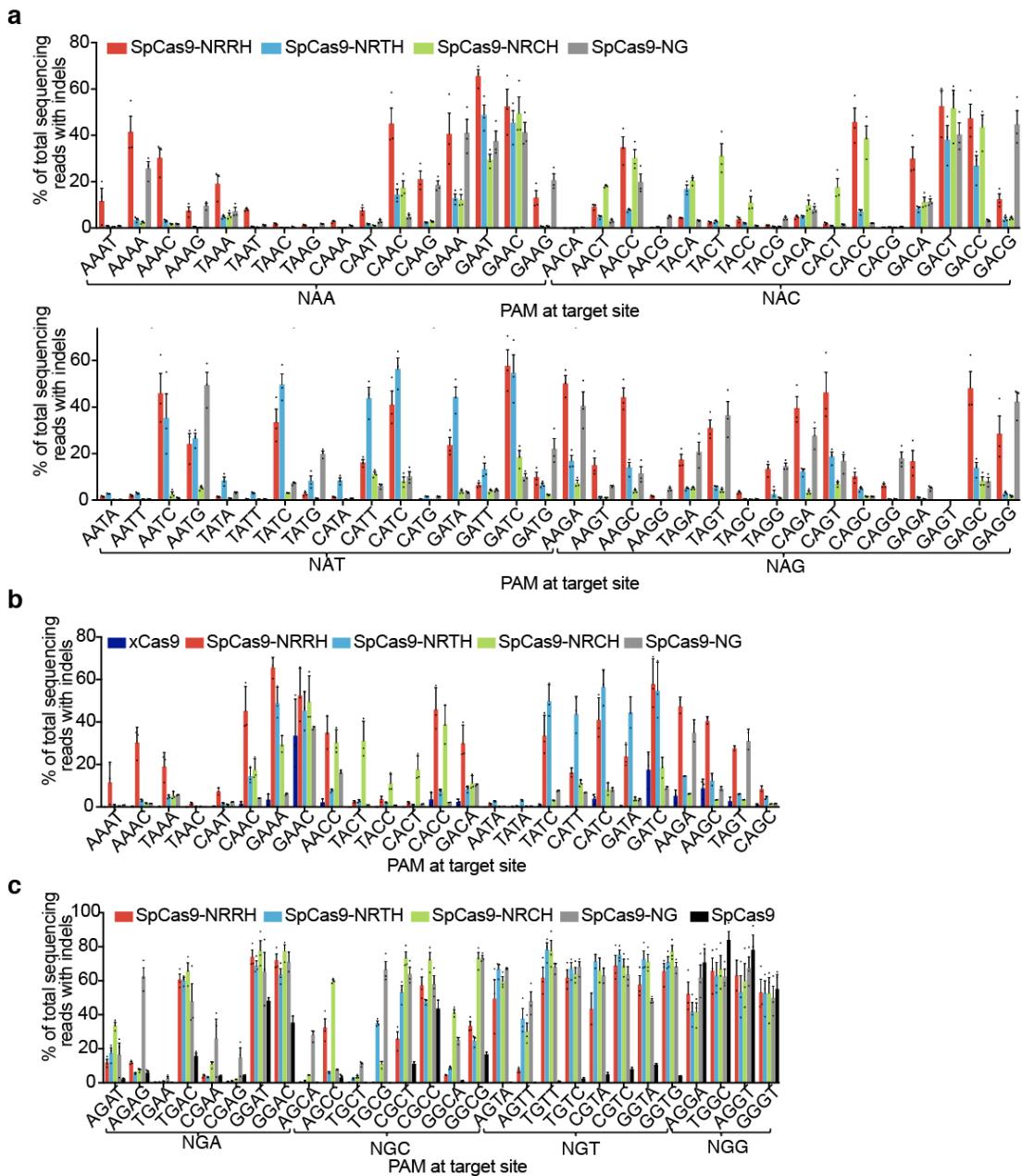
(a) Addition of the Y1131C mutation, which was enriched in the later phases of the NAT evolution trajectory, inactivates BE3-NRTH in HEK293T cells. Mean and SEM of n=3 independent biological replicates are shown. (b) The N-terminal mutations of SpCas9-NRRH, -NRCH, and NRTH mapped to the SpCas9 crystal structure (4UN3). (c) CBE activity of BE3-NRRH, BE3-NRTH, and BE3-NRCH with and without the N-terminal mutations shown in (b) in HEK293T cells. Mean and SEM of n=3 independent biological replicates are shown, with individual values shown as dots.



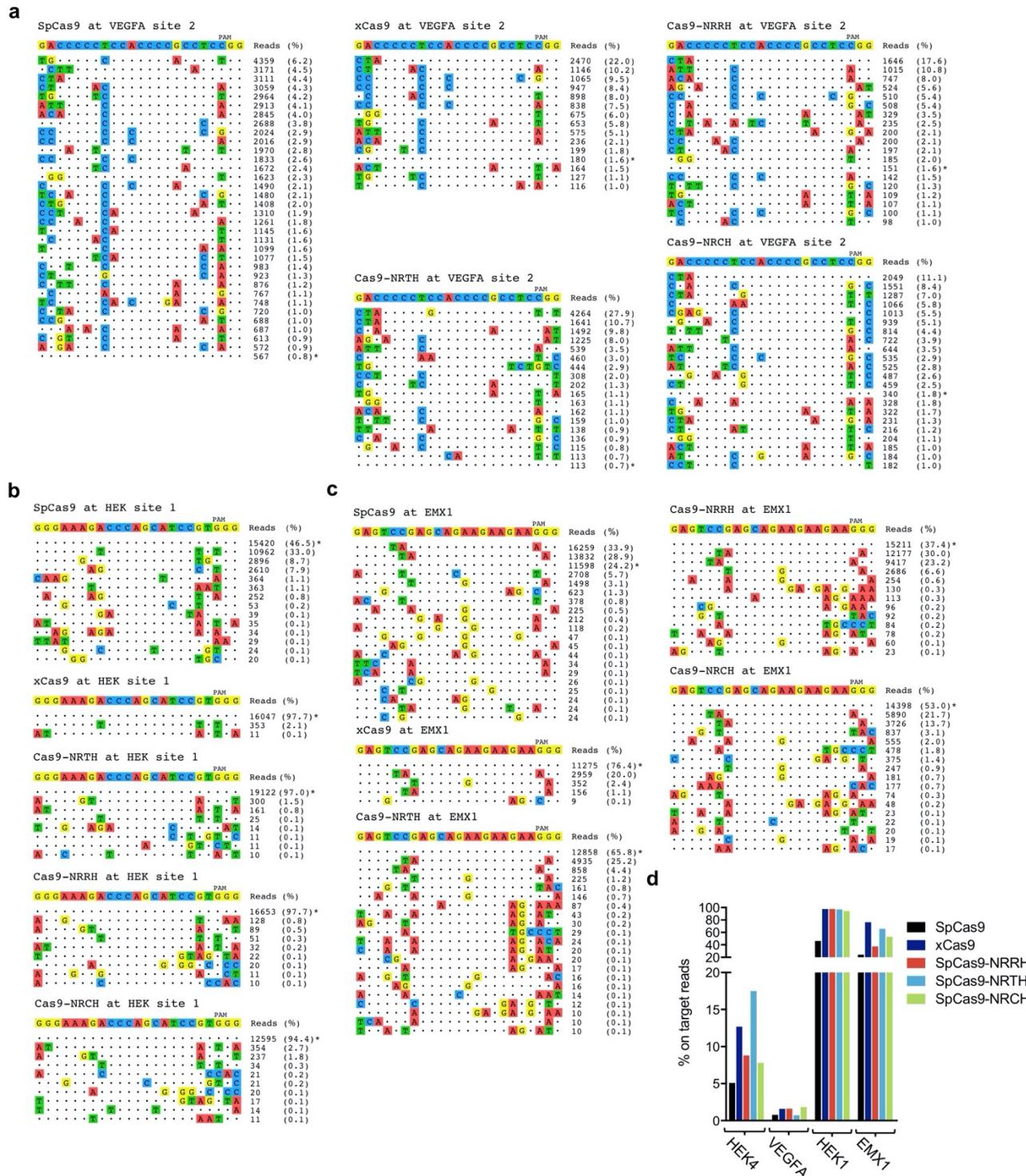
Supplementary Figure 3. Characterization of Cas9 variants evolved in this study, and SpCas9-NG in bacterial PAM depletion experiments. Bacterial PAM depletion of SpCas9-NRRH, -NRCH, -NRTH, and SpCas9-NG on a bacterial NNNN PAM library with 1 h, 3 h, and overnight Cas9 induction. The inverse of the depletion score was used to generate enrichment scores of activity on each NNNN PAM (see Methods for details), which were then used to create sequence logos (WebLogo3.0).



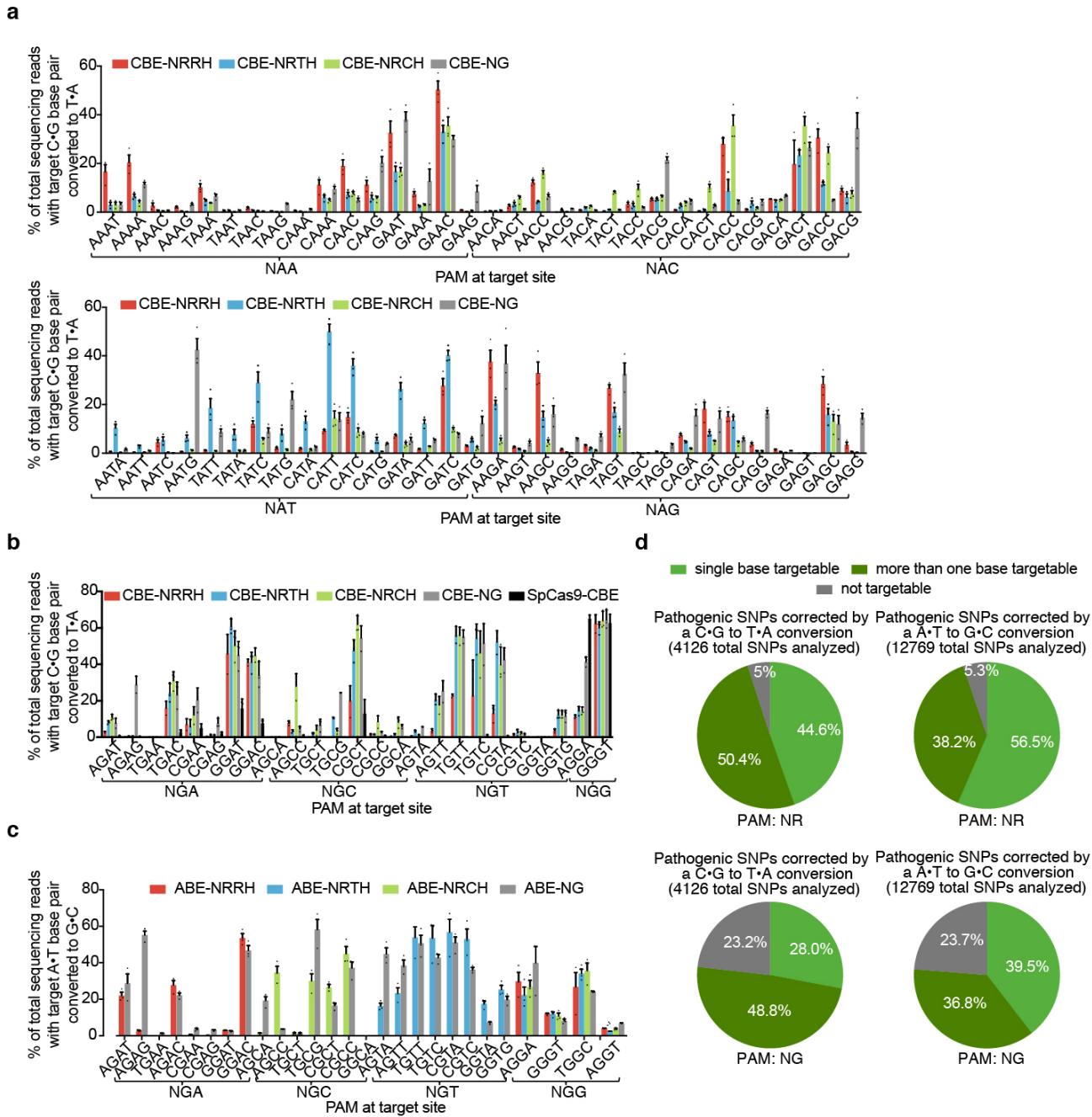
Supplementary Figure 4. Characterization of PAM preferences using a genomically integrated 11,776-member human cell target sequence library. (a) Heat map of base editing activity on the 11,776-member NNNN PAM library in HEK293T cells, with positions 2, 3, and 4 of the PAM defined. For each construct, the mean editing across all sites containing the designated PAM over two independent biological replicates, internally normalized to the highest editing value for each construct, is shown. (b,c) Effect of sgRNA length and 5'G mismatches on the base editing efficiency of profiled SpCas9 variants. The percent decrease of editing efficiency from using a 21 nt sgRNA with either a matched (b) or mismatched (c) 5' G compared to using a matched 20 nt sgRNA is shown for CBE-SpCas9, CBE-NRRH, CBE-NRTH, CBE-NRCH, and CBE-NG on all library sequences containing NAN, NRN, NGN, or NGG PAMs. The mean and SEM are plotted of n=659 (20 nt match NAN), n=456 (21 nt match NAN), n=1922 (21 nt mismatch NAN), n=1328 (20 nt match NRN), n=914 (21 nt match NRN), n=3844 (21 nt mismatch NRN), n=669 (20 nt match NGN), n=458 (21 nt match NGN), n=1922 (21 nt mismatch NGN), n=174 (20 nt match NGG), n=117 (21 nt match NGG), n=480 (21 nt mismatch NGG). (d) Relative editing activities amongst NNNN PAMs in the 11,776-member library for SpCas9-CBE, CBE-NRRH, CBE-NRTH, and CBE-NRCH by nucleotide at the first, second, third, or fourth position of the PAM. (e) Product distribution among edited DNA sequencing reads (reads in which the target C is mutated) within the 11,776-member NNNN PAM library in HEK293T cells. The bars represent mean and SEM for n=11,776 target sites.



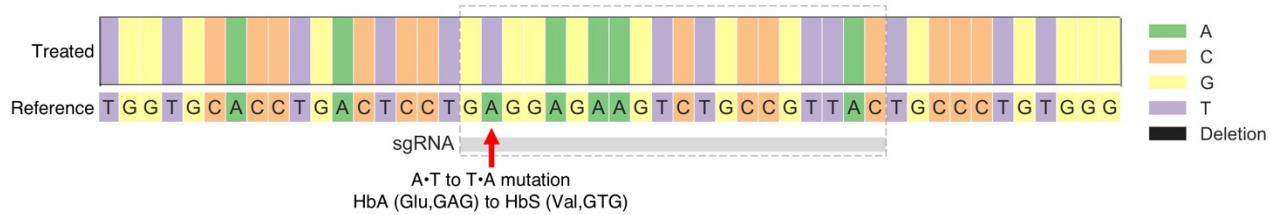
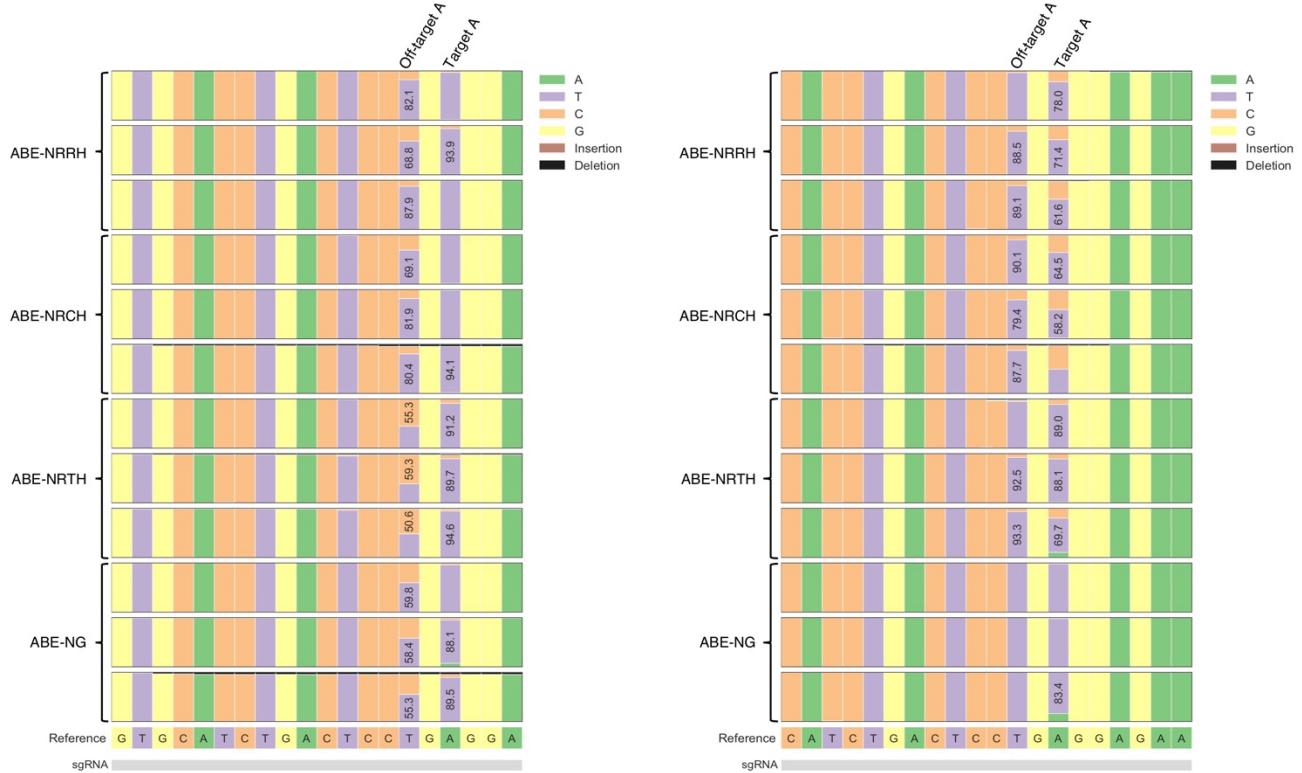
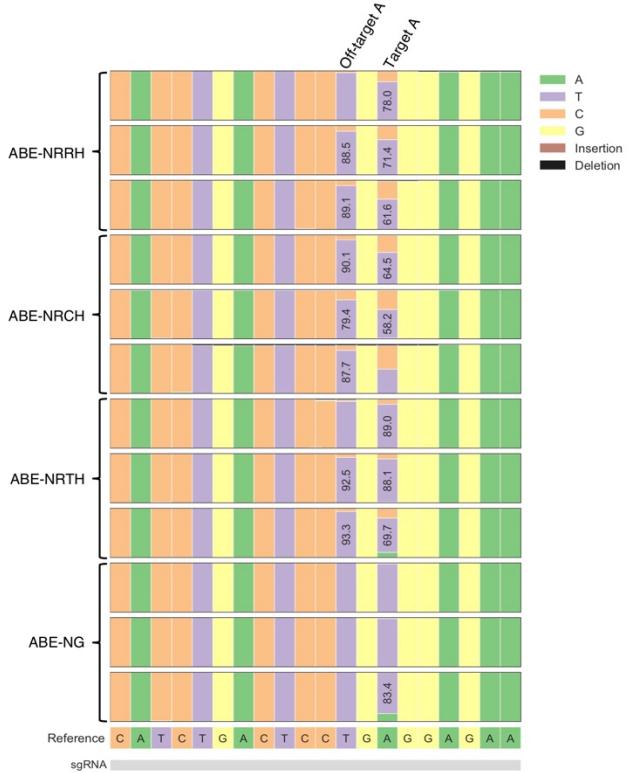
Supplementary Figure 5. Characterization of evolved variants, SpCas9-NG, and SpCas9 in human cell indel formation experiments. (a) Indel formation in HEK293T cells across 64 endogenous human sites containing NANN PAMs for SpCas9-NRRH, -NRTH, -NRCH, and SpCas9-NG. (b) Indel formation in HEK293T cells across endogenous human sites containing NANN PAMs for xCas9, SpCas9-NRRH, -NRTH, -NRCH, and SpCas9-NG. (c) Indel formation in HEK293T cells across endogenous human sites containing NGNN PAMs for SpCas9-NRRH, -NRTH, -NRCH, SpCas9-NG, and SpCas9. Mean and SEM of n=3 independent biological replicates are shown, with individual values shown as dots.



Supplementary Figure 6. Characterization of off-target activity of evolved Cas9 variants. GUIDE-seq analysis of SpCas9, xCas9, and SpCas9-NRRH, SpCas9-NRTH, and SpCas9-NRCH on (a) VEGFA site 2, (b) HEK site 1, and (c) EMX1. On-target reads are denoted by an asterisk. Off-target reads that represent <1% of total reads are omitted for VEGFA site 2, and off-target reads that represent <0.1% of total reads are omitted for HEK site 1 and EMX1. A full list of identified off-target sites is available in Supplementary Table 4.(d) DNA targeting specificity of SpCas9, xCas9, and evolved variants SpCas9-NRRH, -NRTH-, and NRCH resulting from GUIDE-seq analysis on HEK site 4, VEGFA site 2, HEK site 1, and EMX1 in U2OS cells as determined by percentage of on-target reads.

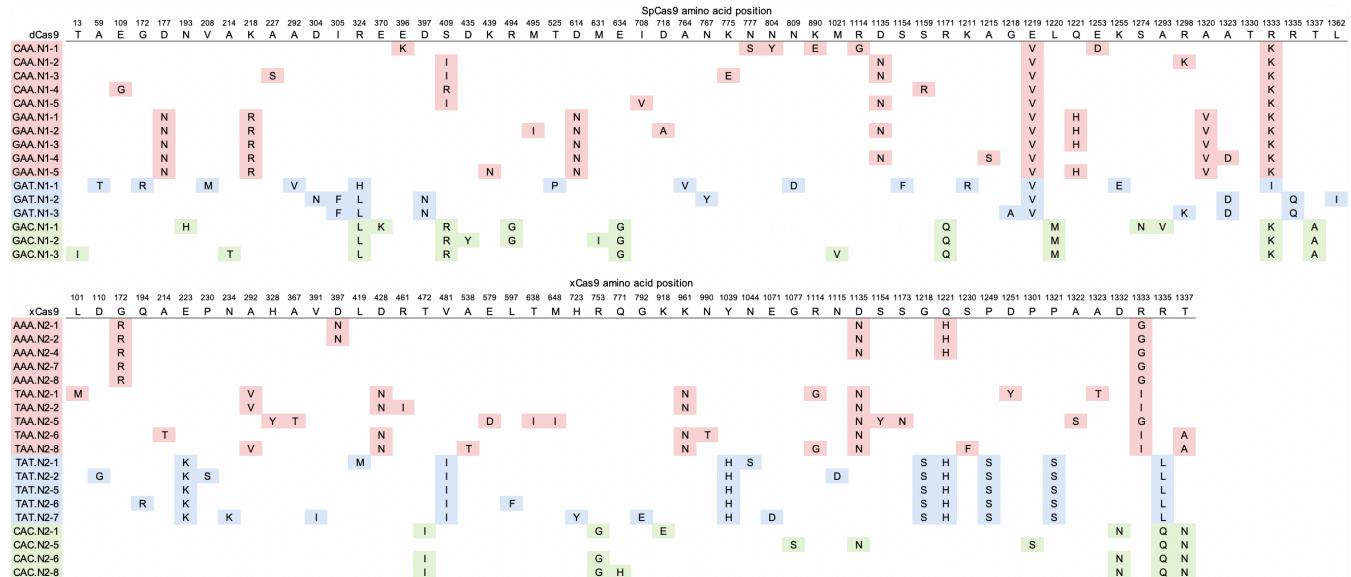


Supplementary Figure 7. Characterization of evolved variants, SpCas9-NG, and SpCas9 in human base editing experiments. (a) Cytosine base editing in HEK293T cells across 64 endogenous human sites containing NANN PAMs for CBE-NRRH, CBE-NRTH, CBE-NRCH, and CBE-NG. (b) CBE in HEK293T cells across endogenous human sites containing NGNN PAMs for CBE-NRRH, CBE-NRTH, CBE-NRCH, and CBE-NG. (c) ABE in HEK293T cells across endogenous human sites containing NGNN PAMs for ABE-NRRH, ABE-NRTH, ABE-NRCH, and ABE-NG. For (a-c), bars represent mean and SEM of n=3 independent biological replicates, with individual values shown as dots. For target sites with NGA, NGC, and NGT PAMs, only ABE-NRRH, ABE-NRTH, and ABE-NRCH are shown, respectively, in addition to SpCas9-NG. (d) Fraction of pathogenic SNPs in the ClinVar Database with either a single targetable base within the window or multiple targetable bases that could in principle be corrected by a C•G-to-T•A (top left) or A•T-to-G•C (top right) base conversion using NR PAMs or C•G-to-T•A (bottom left) or A•T-to-G•C (bottom right) base conversion using NG PAMs.

a**b****c**

Supplementary Figure 8. Sickle-cell anemia allele editing by ABEs derived from evolved SpCas9 variants. (a) Crispresso2 output showing the *HbS* mutation in a engineered HEK293T cell line. HEK293T cells were treated with nickase-SpCas9, sgRNA (binding shown in grey), and ssODN containing the *HbS* point mutation. After two rounds of transfection, sorting, and growth, the cell line sequenced above was isolated and identified to have 100% conversion to the sickle cell anemia allele. (b) Crispresso2 output showing ABE activity of ABE-NRRH, ABE-NRTH, ABE-NRCH, and ABE-NG in *HbS* engineered HEK293T cells using an sgRNA targeting a protospacer (grey bar) preceding a CATG PAM. (c) Crispresso2 output showing ABE activity of ABE-NRRH, ABE-NRTH, ABE-NRCH, and ABE-NG in *HbS* engineered HEK293T cells using an sgRNA targeting a protospacer (grey bar) preceding a CACC PAM.

Supplementary Table 1. PANCE-evolved variant mutation table



Supplementary Table 2. PACE-evolved variant mutation table.

(See the separately provided Excel file.)

Supplementary Table 3. Bacterial PAM depletion scores.

PAM	1 h induction depletion scores				3 h induction depletion scores				ON induction depletion scores			
	SpCas9-NRRH	SpCas9-NRTH	SpCas9-NRCH	SpCas9-NG	SpCas9-NRRH	SpCas9-NRTH	SpCas9-NRCH	SpCas9-NG	SpCas9-NRRH	SpCas9-NRTH	SpCas9-NRCH	SpCas9-NG
AAAA	26	7	6	6	92	12	8	8	1385	14	9	7
AAAC	86	18	21	9	3029	23	31	7	2890	67	162	7
AAAG	29	6	8	15	55	7	7	47	3653	13	8	488
AAAT	46	11	10	8	270	9	14	9	2492	13	19	9
AACA	12	9	16	13	14	10	12	11	28	174	22	10
AACC	21	11	21	6	99	11	39	10	204	7	362	10
AACG	7	7	8	12	8	7	7	31	9	5	9	113
AACT	14	9	24	8	14	12	51	10	24	23	4078	9
AAGA	61	19	14	30	1194	31	17	56	3279	1824	77	2311
AAGC	900	70	20	16	5611	118	66	76	5721	319	5981	1207
AAGG	32	12	14	120	205	15	21	162	7503	10	117	791
AAGT	182	29	17	25	1371	59	31	148	5416	3012	5843	1179
AATA	14	42	9	10	34	79	10	12	2326	1294	27	39
AATC	77	318	14	8	4765	200	15	13	4471	2486	79	15
AATG	12	18	13	56	29	25	9	157	92	145	13	549
AATT	13	48	12	6	19	95	15	8	1887	2098	24	10
ACAA	11	6	6	6	12	7	5	6	17	2107	5	6
ACAC	12	7	9	5	19	10	9	5	100	7	9	4
ACAG	10	7	5	10	10	5	5	9	11	6	6	12
ACAT	9	15	6	6	12	17	7	6	14	80	6	5
ACCA	4	5	4	4	3	5	4	4	3	4	4	3
ACCC	6	4	9	4	5	5	8	4	4	5	12	4
ACCG	3	4	4	5	3	5	4	4	4	4	4	4
ACCT	4	5	10	4	4	6	11	4	3	3	11	4
ACGA	10	8	14	12	10	7	19	18	13	41	37	22
ACGC	18	10	19	11	34	9	37	12	3018	22	118	14
ACGG	9	8	16	16	11	7	16	30	15	8	28	236
ACGT	17	34	26	23	18	82	48	32	65	5384	4752	144
ACTA	4	6	5	4	4	5	5	4	2	10	5	4
ACTC	12	7	7	5	13	8	6	4	333	10	5	4
ACTG	4	5	4	8	3	5	4	9	5	3	4	10
ACTT	5	6	6	4	5	6	6	4	3	3	5	4
AGAA	65	64	48	384	499	188	280	1450	2747	4	3192	1932
AGAC	3880	351	560	71	4834	1821	4063	2341	4934	2744	5152	3119
AGAG	53	63	49	1329	124	67	140	2810	5701	3170	6184	3744
AGAT	259	491	173	1726	5046	1901	1414	489	4938	2746	5378	3256
AGCA	22	34	164	62	61	163	780	197	2332	3889	6919	4189
AGCC	109	65	614	22	1673	84	5622	162	6762	3760	7130	4316
AGCG	11	20	43	852	14	29	332	1196	32	5761	215	6372
AGCT	29	158	791	63	145	262	834	211	212	3878	7400	4479
AGGA	297	169	142	9474	3475	873	730	673	6778	3769	7407	408
AGGC	1723	297	559	1253	9079	489	694	733	9856	5480	9678	5858
AGGG	118	53	124	9048	1007	99	847	1463	12945	7198	12879	2599
AGGT	265	725	595	15281	10838	1021	9108	750	10932	6079	11552	1749
AGTA	116	1038	418	486	602	2040	4551	874	5215	2900	5772	3494
AGTC	294	7613	360	397	7627	2873	1603	528	7657	4258	8130	4921
AGTG	39	458	307	14785	122	1848	1178	951	10577	5881	5229	6330
AGTT	104	1272	337	1073	358	3096	6907	995	7676	4269	8759	5302
ATAA	8	5	4	5	8	6	4	6	5	7	4	5
ATAC	24	9	8	6	46	9	6	5	4764	10	6	5
ATAG	7	6	5	13	7	6	4	18	8	3	4	66
ATAT	10	11	5	5	12	10	5	6	24	8	5	5
ATCA	5	6	4	5	4	6	4	5	3	8	4	4
ATCC	7	7	6	5	7	5	6	5	5	4	4	5
ATCG	4	5	4	7	4	4	4	7	3	7	4	6
ATCT	4	6	6	4	4	5	6	5	4	4	7	4
ATGA	13	11	11	19	14	10	13	40	67	39	18	113
ATGC	46	11	16	14	134	12	17	19	10195	7	24	62
ATGG	9	9	12	23	9	7	14	81	15	5	9	681
ATGT	17	25	16	22	39	32	27	57	1696	300	417	757
ATTA	4	5	4	4	4	5	5	4	3	3	4	4
ATTC	12	7	4	4	15	6	5	4	83	18	6	5
ATTG	4	5	4	12	3	5	4	13	3	4	4	18
ATTT	4	6	4	3	4	5	5	4	3	6	4	4
CAAA	83	11	16	11	724	10	32	17	3800	39	2314	49

CAAC	998	31	68	13	7012	133	185	16	7611	4232	7474	86
CAAG	96	11	17	29	354	12	14	134	1818	10	60	635
CAAT	681	30	22	13	6980	59	77	19	6928	3852	3721	108
CACA	23	17	24	16	45	24	80	27	728	4453	7687	582
CACC	76	17	79	15	357	31	195	18	9489	5277	9113	75
CACG	14	8	15	29	18	9	17	59	55	7	66	409
CACT	23	18	109	15	71	25	395	20	9690	5388	9995	164
CAGA	180	26	16	18	1057	34	26	116	7896	4391	1972	2387
CAGC	878	90	33	21	11885	214	61	79	13393	7447	12669	1279
CAGG	83	13	16	72	276	14	18	171	15175	49	79	1135
CAGT	294	60	21	43	11492	155	78	242	12482	6941	12249	7415
CATA	34	414	25	17	133	2650	30	21	7067	3930	188	163
CATC	491	983	32	13	11143	4197	142	19	11854	6591	2376	360
CATG	29	62	21	67	68	186	26	280	14700	8174	156	1070
CATT	41	531	35	14	182	4176	142	18	11210	6233	3939	77
CCAA	5	4	4	4	4	5	4	4	4	10	4	4
CCAC	11	5	5	5	10	5	6	4	12	6	4	3
CCAG	4	4	4	5	4	4	4	4	4	4	4	4
CCAT	7	5	6	5	7	6	5	4	4	5	5	4
CCCA	3	4	4	4	3	4	4	4	3	3	3	4
CCCC	4	4	9	4	4	4	9	4	4	3	8	3
CCCG	3	4	4	5	3	4	4	4	5	5	4	4
CCCT	4	5	13	4	3	5	16	4	3	6	14	4
CCGA	5	5	6	5	4	5	6	5	4	4	5	4
CCGC	13	6	7	5	12	6	7	4	19	4	6	4
CCGG	4	4	5	6	3	4	4	5	4	6	5	5
CCGT	8	7	7	7	6	6	6	5	4	7	5	5
CCTA	4	6	5	4	4	6	5	5	3	8	4	4
CCTC	11	8	8	4	11	8	7	3	29	24	5	4
CCTG	3	5	4	6	3	5	4	5	3	8	4	5
CCTT	5	6	8	4	4	6	8	4	3	4	8	4
CGAA	124	117	172	181	619	103	953	144	6697	3724	7252	4390
CGAC	1746	765	1359	66	4449	1676	7478	154	9991	5555	9484	5741
CGAG	55	67	188	8479	305	93	379	5453	12131	6746	12002	7265
CGAT	219	1618	1033	1062	5472	4122	4599	589	11387	6332	11664	7060
CGCA	30	105	472	103	85	268	3382	209	13871	7713	12866	557
CGCC	147	133	5673	30	3973	409	10016	152	6257	6958	12703	7689
CGCG	15	35	118	1203	15	75	479	1143	28	10523	8804	2665
CGCT	41	431	1289	122	168	1022	11402	200	15159	8429	14461	8754
CGGA	127	204	283	1327	979	218	535	1027	13281	7385	13556	8206
CGGC	1358	463	2055	507	15825	1491	6650	384	18129	10081	16868	10211
CGGG	75	79	449	2649	240	97	3184	656	11368	12642	20187	12220
CGGT	294	491	1314	2823	2108	1021	3985	1021	22214	3088	20214	12236
CGTA	212	11751	643	1836	1613	4251	1898	1822	11819	6572	6015	7282
CGTC	3157	5322	1560	832	14330	5398	12043	992	16058	8929	15274	9246
CGTG	69	1240	398	26135	113	2982	13306	1917	2337	10396	8438	10215
CGTT	141	17891	12236	2097	5558	6280	14012	8074	17994	10006	17771	10757
CTAA	11	6	5	5	17	6	5	5	10	4	5	5
CTAC	27	10	9	5	82	7	11	4	851	10	10	4
CTAG	11	5	4	10	10	5	5	12	8	6	4	14
CTAT	18	9	7	5	40	7	7	5	1112	3	7	5
CTCA	5	6	6	5	5	4	6	4	3	4	5	4
CTCC	14	7	15	5	14	6	12	4	115	8	32	3
CTCG	4	4	4	7	4	5	5	6	4	4	4	5
CTCT	7	7	15	5	6	7	14	5	3	10	23	5
CTGA	13	7	7	11	14	6	6	11	74	3	6	11
CTGC	53	9	10	9	163	9	8	7	20035	6	8	6
CTGG	8	5	5	15	7	5	5	25	6	6	4	75
CTGT	24	10	10	12	50	11	9	13	6900	8	7	31
CTTA	5	7	6	5	5	7	5	5	4	3	5	4
CTTC	26	12	10	4	68	13	8	5	2714	39	6	4
CTTG	4	5	4	10	4	5	5	10	4	5	4	9
CTTT	6	9	7	4	6	11	7	4	6	18	6	4
GAAA	378	39	27	19	2983	41	68	81	2882	1603	3180	963
GAAC	1241	295	78	18	4743	894	307	115	4733	2632	2528	1020
GAAG	218	27	21	213	1195	41	40	964	5937	3301	910	3854
GAAT	1123	128	46	17	4870	84	241	148	263	2778	5191	3143
GACA	72	35	34	20	237	53	78	100	6511	3621	6548	3964
GACC	141	43	212	19	679	144	1284	42	6697	3724	6512	180
GACG	27	15	19	54	63	21	30	125	4232	48	729	1764
GACT	85	49	107	17	285	61	1037	77	7094	3945	7890	2388

GAGA	500	69	16	152	5224	328	47	362	4764	2649	5568	1124
GAGC	2215	1050	56	130	8318	3133	195	403	8448	4697	8866	5367
GAGG	327	37	18	444	3342	62	56	971	10789	5999	2671	3234
GAGT	1698	312	54	453	9369	1765	95	1135	9717	5403	9987	6045
GATA	202	364	25	22	852	1924	120	104	1708	2849	5444	1099
GATC	417	527	38	17	7804	2940	153	67	7942	4417	8318	5035
GATG	135	166	24	570	198	292	70	697	11009	6122	2682	3247
GATT	139	788	70	19	2148	3236	314	58	7923	4406	9157	1848
GCAA	13	5	6	6	17	5	6	6	131	8	5	5
GCAC	40	10	13	7	97	9	10	5	11518	5	12	4
GCAG	10	6	5	11	9	5	5	12	9	6	3	16
GCAT	21	8	11	7	59	8	10	7	1640	6	9	7
GCCA	5	6	8	6	4	5	7	6	4	8	6	6
GCCC	11	7	19	5	12	7	31	5	37	6	118	4
GCCG	4	5	5	9	3	5	5	8	3	3	6	8
GCCT	6	7	26	6	6	68	6	5	6	2845	5	
GCGA	14	7	11	10	18	7	9	10	115	9	8	10
GCGC	49	14	13	9	181	14	11	9	13751	16	13	8
GCGG	8	5	7	13	7	4	7	12	6	4	6	20
GCGT	22	13	16	13	76	10	12	16	4334	12	18	49
GCTA	8	8	11	6	8	7	9	6	7	5	11	6
GCTC	26	20	15	6	224	28	17	6	5011	5573	51	6
GCTG	5	5	6	12	5	5	5	15	4	9	4	17
GCTT	11	12	14	6	14	12	21	6	88	8	256	5
GGAA	359	143	365	957	6436	221	5409	624	6160	3426	6860	4153
GGAC	16350	862	3535	162	9403	1771	7902	176	10396	5781	10022	6067
GGAG	219	225	708	4846	10237	276	1721	4958	10400	5783	10912	6605
GGAT	1212	1787	734	520	10360	3902	8707	419	10781	5995	11043	6684
GGCA	76	148	226	170	358	305	1458	281	13609	7567	12944	7835
GGCC	541	409	3460	38	11882	236	4993	131	490	7074	12665	7667
GGCG	24	52	151	1371	64	98	255	3374	3923	1745	1486	4495
GGCT	105	397	3024	197	3643	2744	12244	321	15564	8655	15529	9400
GGGA	697	457	582	4336	11900	747	3334	412	12405	6898	12684	240
GGGC	3333	1124	1048	1394	2610	2949	13158	690	16953	9427	8344	10102
GGGG	325	206	508	3878	17489	169	2450	2823	19418	5399	9321	11285
GGGT	742	983	3528	3626	19273	3630	8099	4667	20748	11537	20543	12435
GGTA	321	2683	1468	795	10583	3986	8894	1709	10793	667	11280	6828
GGTC	2428	5115	5247	1349	14636	5513	6150	1013	15433	8582	15600	9443
GGTG	276	1322	1938	3983	2891	6532	1620	560	19942	71	18486	11190
GGTT	503	7943	10864	677	16075	6055	13510	1557	2283	8884	17135	10372
GTAA	22	11	8	9	76	11	7	10	5775	16	8	16
GTAC	98	23	19	10	3106	31	13	12	9659	5371	57	20
GTAG	17	9	9	24	31	9	6	123	305	6	5	381
GTAT	44	21	13	13	117	27	11	16	10488	71	23	64
GTCA	8	9	8	7	9	10	8	7	8	7	7	7
GTCC	18	11	17	8	58	15	21	6	1300	9	68	6
GTCG	6	8	6	13	5	7	5	15	4	5	5	33
GTCT	11	12	14	8	13	15	14	8	39	10	46	7
GTGA	40	15	11	18	98	14	11	49	10430	19	13	521
GTGC	282	22	14	16	10282	21	15	27	12004	6675	49	474
GTGG	19	8	8	47	34	8	7	350	1130	8	6	259
GTGT	104	21	15	20	464	29	15	107	20043	1115	56	584
GTTA	12	10	8	7	16	9	7	8	36	10	6	6
GTTC	61	20	15	7	367	33	12	7	1329	903	19	5
GTTG	8	9	8	24	7	8	6	56	5	6	5	272
GTTT	12	14	11	6	17	20	9	7	44	56	12	6
TAAA	26	8	9	9	109	9	9	10	704	13	9	9
TAAC	268	21	21	10	2537	60	62	12	5111	2842	1082	15
TAAG	21	9	8	20	65	8	9	54	2401	7	9	299
TAAT	111	19	16	10	404	20	17	13	5192	71	233	31
TACA	10	11	16	11	14	11	24	12	12	9	178	16
TACC	25	12	31	10	92	11	197	9	431	16	2243	8
TACG	7	8	12	16	6	7	10	29	6	6	10	104
TACT	13	12	42	10	14	13	117	11	43	19	7654	15
TAGA	39	16	15	16	140	19	13	86	5401	112	19	244
TAGC	391	64	16	14	1655	163	37	27	10916	1214	1176	801
TAGG	26	11	12	54	63	10	10	155	12845	7	14	651
TAGT	69	30	19	22	300	45	18	97	10589	144	151	469
TATA	13	145	16	10	23	229	14	13	82	2433	130	21
TATC	247	562	21	12	1016	3061	36	12	8467	4708	8663	18
TATG	14	39	12	33	14	52	12	63	51	6801	15	607

TATT	17	262	19	8	29	631	31	11	849	4245	446	14
TCAA	5	4	4	4	4	4	4	5	3	7	4	5
TCAC	8	5	5	4	7	6	5	4	11	6	5	4
TCAG	5	5	4	6	4	4	4	5	3	5	3	5
TCAT	6	6	5	4	6	5	5	5	4	4	4	5
TCCA	3	4	4	4	3	5	4	4	4	4	4	4
TCCC	4	5	6	4	4	4	5	4	4	6	5	4
TCCG	3	4	4	5	3	4	5	4	3	4	4	4
TCCT	3	4	7	4	3	5	6	4	4	4	6	3
TCGA	5	5	10	7	4	6	8	6	4	6	7	6
TCGC	11	7	9	6	9	5	11	5	12	6	9	5
TCGG	5	5	8	10	4	5	6	7	4	6	6	7
TCGT	9	9	13	9	6	8	10	8	5	9	17	8
TCTA	4	5	5	4	3	5	4	4	3	3	4	3
TCTC	10	7	4	4	9	7	5	4	10	6	5	4
TCTG	3	5	4	6	3	4	4	6	3	5	4	6
TCTT	4	6	6	4	4	5	6	4	3	4	6	4
TGAA	30	284	285	905	131	257	1431	1099	7121	283	7255	4392
TGAC	333	1288	2349	137	9275	3494	3898	562	10361	5761	9887	5985
TGAG	28	167	105	6079	76	376	348	5799	4349	7254	12764	7726
TGAT	71	2381	543	558	636	3590	8009	1539	9578	5326	10158	6149
TGCA	17	282	472	169	22	431	2642	677	181	7711	13404	2029
TGCC	45	292	2172	58	252	2321	2590	299	14372	7992	657	7951
TGCG	9	73	70	1166	8	98	129	1086	7	12983	21501	2603
TGCT	18	948	566	472	61	1108	12361	310	201	9011	15677	9489
TGGA	41	215	133	9186	183	321	290	3091	13142	7308	1944	4118
TGGC	338	556	512	2578	814	2913	12998	375	18438	10252	16485	3327
TGGG	27	145	200	18433	77	225	454	3660	8791	14664	4834	7315
TGGT	104	784	797	7985	638	1282	8580	989	22847	12704	21762	6587
TGTA	47	5471	499	1710	134	2087	1862	5365	1223	6119	11808	7148
TGTC	439	8863	1866	1662	15211	5729	12784	614	17829	9914	16213	9814
TGTG	30	1310	672	11045	73	7508	1862	9652	11852	6590	21246	12861
TGTT	58	6345	1085	2676	274	6693	4978	1721	19144	10645	18939	5732
TTAA	4	4	3	4	4	5	4	4	3	8	4	4
TTAC	12	5	4	4	10	5	5	4	17	5	4	4
TTAG	4	4	3	7	4	5	4	6	3	5	5	6
TTAT	6	5	4	4	5	5	4	4	4	5	4	4
TTCA	3	4	4	4	3	5	4	4	3	12	4	4
TTCC	5	5	5	4	5	5	5	4	4	14	4	4
TTCG	3	4	4	5	3	4	4	4	5	5	4	4
TTCT	3	5	6	4	3	5	6	4	3	5	5	4
TTGA	5	5	6	8	5	5	5	7	4	5	5	6
TTGC	19	6	7	7	26	5	6	5	126	4	7	4
TTGG	4	5	5	12	4	5	5	11	4	5	4	16
TTGT	9	7	7	9	9	8	6	10	12	7	5	7
TTTA	3	4	4	4	3	5	4	4	3	2	4	3
TTTC	9	5	4	4	8	5	5	4	8	8	5	4
TTTG	3	4	4	5	3	5	4	5	3	4	4	4
TTTT	3	5	4	3	3	5	4	4	3	5	5	4

Supplementary Table 4. GUIDE-seq identified off-target sites

(See the separately provided Excel file.)

Supplementary Table 5. Genomically integrated 11,776-member human cell target sequence library editing data.

(See the separately provided Excel file.)

Supplementary Table 6. Sample sizes (n) in Figure 3b.

	NG	BE4	ax	fn	es
NAA	744	750	666	749	727
NAC	746	748	695	747	731
NAG	740	745	674	746	716
NAT	742	747	672	741	724
NGA	749	749	684	743	735
NGC	747	746	687	753	737
NGG	763	760	705	762	747
NGT	752	747	690	747	738

Supplementary Table 7. Plasmids and selection phage (SP) used in this work

Name	Class (res)	Origin	ORF1 (prom [RBS] genes)	ORF2 (prom [RBS] genes)	ORF2 (prom [RBS] genes)
pSM072a	AP (carb ^R)	SC101	(G7' protospacer +PAM) P _{lac} [sd8] ¹ gIII, luxAB	P _{lac} G7' sgRNA	
pTW168a	AP (spec ^R)	ColE1	(G7' protospacer +PAM) P _{lac} [sd8] gIII-C	P _{lac} G7' sgRNA	
pTW169a	AP (carb ^R)	SC101	(Doench ON2 protospacer +PAM) P _{lac} [sd8] gIII-N	P _{lac} Doench ON2 sgRNA	
pTW199b	AP (spec ^R)	ColE1	(G7' protospacer +PAM) P _{lac} [sd2] gIII, luxAB	P _{lac} G7' sgRNA	
pTW170b	AP (kan ^R)	p15A	(VEGF2 protospacer +PAM) P _{lac} [SD8] gVI	P _{lac} s13 sgRNA	
pTW099c	AP (carb ^R)	SC101	P _{dsp} [SD8] gIII, gVI		
pTW221b1	CP (carb ^R)	SC101	P _{pro1} ² [sd8] rpoZ(I12N)-dSpCas9(1-573 A10T/V322I/S409I/E427G)		
pTW221b3	CP (carb ^R)	SC101	P _{pro3} [sd8] rpoZ(I12N)-dSpCas9(1-573 A10T/V322I/S409I/E427G)		
pTW221b5	CP (carb ^R)	SC101	P _{pro5} [sd8] rpoZ(I12N)-dSpCas9(1-573 A10T/V322I/S409I/E427G)		
pTPH308	Selection plasmid for PAM depletion (kan ^R)	pUC	HEK3 protospacer + NNNNN PAM	P _{pro1} [SD8] sfGFP	
pTW218	Selection plasmid for bacterial Cas9 nuclease selection (kan ^R)	pUC	HEK3 protospacer + target PAM	P _{pro1} [sd8] B. subtilis sacB levansucrase	P _{pro1} [SD8] sfGFP
pTW222	Bacterial Cas9 expression plasmid for cutting selections (carb ^R)	SC101	pBAD [sd8] Cas9	P _{lac} HEK3 sgRNA	
pSM060	Mammalian expression (carb ^R)	pUC	P _{CMV} BE3		
pSM063	Mammalian expression (carb ^R)	pUC	P _{hU6} sgRNA		
pSM067	Mammalian expression (carb ^R)	pUC	P _{hU6} sgRNA		
pSM068	Mammalian expression (carb ^R)	pUC	P _{CMV} BE4max		
pSM069	Mammalian expression (carb ^R)	pUC	P _{CMV} ABEmax		
pSM070	Mammalian expression for Tol2 mediated integration (carb ^R)	pUC	P _{CMV} BE4max	P _{SV40} blast ^R	
pSM071	Mammalian expression (carb ^R)	pUC	P _{CMV} Cas9max		
pSM090	mRNA expression (carb ^R)	pUC	P _{T7} Cas9max-polyA		
SP41	ΔgIII SP	M13 f1	P _{gIII} [sd8] rpoZ-dSpCas9		
SP47	ΔgIII SP, recoded gVI promoter	M13 f1	P _{gIII} [sd8] rpoZ-dSpCas9		
SP56	ΔgIII SP, recoded gVI promoter	M13 f1	P _{gIII} [sd8] npuC-dSpCas9(574-1368)		
SP58	ΔgIIIΔgVI SP, recoded gVI promoter	M13 f1	P _{gIII} [sd8] npuC-dSpCas9(574-1368)		

Supplementary Table 8. Primers used in this work.

amplicon	fwd (5'-3')	rev (5'-3')	purpose
SP backbone	CACCGTTCATCTGCCTCTT	CGACCTGCTCCATGTTACTTAG	qPCR estimation of SP titer
SP gene insert	TAATGGAAACTTCCTCATGAAAAAGTC TTTAG	ACAGAGAGATAAACATAAAAACAGGGAAGC	amplification of evolving protein on SP for Sanger sequencing
SP gene insert	GCCAGTTACAAAATAAACAGCC	CACCGTTCATCTGCCTCTT	amplification of evolving protein on SP for DNA shuffling
SpCas9c variant	GAECTCCCTGCAAGCCTCAG	ACAGAGAGATAAACATAAAAACAGGGAAGC	uracil incorporation PCR for DNA shuffling
SpCas9c variant	ATGTTGAAAATCTCCTCUAGATCACT CTCCACCGAGCTGAG	ATGCCAGCAUTGTTGACTCTGTTGAAATCAG	DNA shuffling
PAM depletion library	ACACTCTTCCCTACACGACGCTCTTC CGATCTNNNCAATACGCAAACCGCC TCTC	TGGAGTTCAGACGTGTGCTCTCCGATCTTGCTG TAAGCGGATGC	HTS of PAM depletion
oligo library	TTTTGTTTCTGTGTTCCGTTGTCG TGCTGTAACGAAAGGATGGGTGCGAC GCGTCAT	GTTGATAACGGACTAGCCTTATTAAACTTGCTATGCT GTTCCAGCATAGCTCTAAC	cloning of human 11,7876 integrated target site library
donor template	GTTTAAGAGCTATGCTGGAAACAGC	ACTGCACCGCGTCGCACCCATCCTTCGTTACAGCAC GGACAACGGAACACAGACAAAACAAAAAGCACCGAC TC	cloning of human 11,7876 integrated target site library
plasmid insert	TAACCTGAAAGTATTCGATTTCTGG CTTTATATATCTTGTGGAAAGGACGAA ACACCG	TTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATG TCTGCTCGAACGGGCCGTACCTCTAGAACACTCTTC CCTACACGAC	cloning of human 11,7876 integrated target site library
library PCR1	AATGATAACGGCGACCACCGAGATCTA CAC NNNNNNNN ACACTCTTCCCTACACGAC	GTGACTGGAGTTCAGACGTGTGCTCTTC CGATCT GTGGAAAGGACGAAACACCG	HTS of human 11,7876 integrated target site library
library PCR2	AATGATAACGGCGACCACCGAGATCTA CAC	CAAGCAGAAGACGGCATACGAGAT NNNNNNNN GTGACTGGAGTTCAGACGTGTGCTCTTC	HTS of human 11,7876 integrated target site library
pSM067a1	CACAAAGATTAAAAACCCCCAA	TGCAACCTCACTATGGTATGCT	HTS of genomic target site
pSM067a2r	GTGTCAAATTGAACGGTCTGA	AAAGACAGGCCATCAAAGGAA	HTS of genomic target site
pSM067a3	TTAATGAATTACCCTGACCCC	CTCCAATCTCCCTTTGTTTG	HTS of genomic target site
pSM067a4	CATTCGGAGATTGGATGTTCT	GCATTTCCACAGCTACACCATA	HTS of genomic target site
pSM067a5	CTGCACATCTCGTCATAGTTC	TGGGAAGGTCGCTATTAGAGTG	HTS of genomic target site
pSM067a6r	TGAACAAATCAAGCCTGATGAC	GGTAAGGTACGAAGACCTCTG	HTS of genomic target site

pSM067a7	CTTGGCAATTCTGGTGAAGAT	ATTAAGGTTTGCTATCGGCA	HTS of genomic target site
pSM067a8	CACGAAGGCATATATTTGGTGA	AAAGTTCTAAGGCCCTGGAGA	HTS of genomic target site
pSM067a9r	GTTCCCTGTAAGCCTGGACTC	TGAGGATCTTCATGGCGTAGTA	HTS of genomic target site
pSM067a10	GCCCCAGAGAGATGACAGATAG	TAGGGAAGAAAATATCCTGCG	HTS of genomic target site
pSM067a11	AGATGAAGAGACAAGCTGGGAG	GCAGTACAAGGCCACAGACTTT	HTS of genomic target site
pSM067a12	AGAGAGAGCAGGACGTCACAGT	AGCACTACCTACGTCAGCACCT	HTS of genomic target site
pSM067a13	GCAAAACATACTGACTGGCTGA	ATCAAGAGCAACCTGGACAGAG	HTS of genomic target site
pSM067a14r	CTTCATGATGGCAAGAACATACA	ATTCACTGCAGAGTGCTTCA	HTS of genomic target site
pSM067a15	CTAGAGACAGCCCCAGGAGTC	ATCCTCAAGAAGGAAGTCATCG	HTS of genomic target site
pSM067a16r	AGGCGACCTCTACAGCTTACTG	GAATCTCAGGGACTTTCATGG	HTS of genomic target site
pSM067a17	GGAATCCAGAGTGAGCTTCA	ATTGCATACATTGAAAGACCC	HTS of genomic target site
pSM067a18	CAGGTGGTAACGAGCTGCAT	AAGCACTACCTACGTCAGCACC	HTS of genomic target site
pSM067a19	CTTCAGGAAGTTGTTCTGAGGC	TGAGGAGACACTCCAAGAGAGC	HTS of genomic target site
pSM067a20	CAGAGTCTCTAGGCAGAGGGAA	TAACTCCTCTTTGTTGGGG	HTS of genomic target site
pSM067a21r	ATGGAACAGATCACACATCTGC	AGCTGTCCATTGAGGTACCAAGT	HTS of genomic target site
pSM067a22	AGAGCGAGACTCCATCTCAAAA	CTATCTGGAGTTGGCCTGTT	HTS of genomic target site
pSM067a23	CTGTGCCCTCTTCTCTCAA	TGCAAGAACTATGACGGAGATG	HTS of genomic target site
pSM067a24	GAGGGGTGTGTGAGCGAG	GAGGGAGACGAGCTAAC	HTS of genomic target site
pSM067a25	AAACCCTGTTGCGTTACATT	AAATGGATTGGATCTTCCACAC	HTS of genomic target site
pSM067a26	CTGACTCATGAGGGGACTTTA	CAAGACCTCATCAGGTGAGCAT	HTS of genomic target site
pSM067a27	CTCAGAGACACGGCCTAGTG	CAGAACAAATCCGATTACCGTAG	HTS of genomic target site
pSM067a28	GAGAGAGCAGGACGTCACAGT	CGGTCTCAAGCACTACCTACG	HTS of genomic target site
pSM067a29	CACAAAGATTAAAAACCCCCAA	TGCAACCTCACTATGGTATGCT	HTS of genomic target site
pSM067a30r	CCTGGCTCTCACTCTGACTTT	AGGGTATGCTTCTGCAATCAGT	HTS of genomic target site

pSM067a31r	CCTGGCTCTCACTCTTGACTTT	AGGGTATGCTCTGCAATCAGT	HTS of genomic target site
pSM067a32	CAATGAGTCCTGCTCCACTAGA	GACCTTTTACTCCCAGGCAC	HTS of genomic target site
pSM067a33	TTTCTGGGTTAGGAGTGAAGGA	AGCTGGAGACAGAATACTTGGC	HTS of genomic target site
pSM067a34r	AGACAATTGTTCGACTGCTCA	ATCCTGGAGGTGTTATGTG	HTS of genomic target site
pSM067a35	AGTGATTTGCCCTACCAGTCCT	TGGCAGTCAAACCTCTCT	HTS of genomic target site
pSM067a36	TTACCTCTTAGCACCCCTTCG	TGCCTTTCCAATCAATCTCTT	HTS of genomic target site
pSM067a37	TTCAGAGGATAGCAACATACTCG	AATTGGGCTGATTAAAAGCAT	HTS of genomic target site
pSM067a38	CTACTACCCCCATCTCTCCCTC	CATTAGCTTCCTTGCTTCTC	HTS of genomic target site
pSM067a39	CAACATTCTACAAATGGCCT	ATATATGCCTCGTGTGCCAA	HTS of genomic target site
pSM067a40	CACAGCAGACAATTGTGAAGA	AAAAGCATATGTGTGTCCCC	HTS of genomic target site
pSM067a41	CCGGTACACCACGTTCTCT	ACTTGTTCTGCTGAGTTAGGG	HTS of genomic target site
pSM067a42	ACTGAAGTACCCAGAAGTGGCT	CAGTCCAGATAGTTGTCTCCCC	HTS of genomic target site
pSM067a43	CAATGAGTCCTGCTCCACTAGA	GACCTTTTACTCCCAGGCAC	HTS of genomic target site
pSM067a44	AGGCCTTGCTATGGTAGAACTC	ACCAGCAATTCTTTTCTTG	HTS of genomic target site
pSM067a45	GCTTAAACATTGTCTGTGCG	GTTTCTGCCCCCTCCCTCAGTA	HTS of genomic target site
pSM067a46	CCCTGGGCATTTACTTAATCT	TTCTACCATAGCAAGGCCTAATG	HTS of genomic target site
pSM067a47	ATTGAAGAATTGGAGGGAAGG	TCACAGTTCTGGTACTGTCC	HTS of genomic target site
pSM067a48	CTTGTGATCCAAAAAGTGTCCA	CCTGTTACACGTTCATGTGCT	HTS of genomic target site
pSM067a49	CAATGAGTCCTGCTCCACTAGA	GACCTTTTACTCCCAGGCAC	HTS of genomic target site
pSM067a50	GAGAGTGAGATTCCGTCTAAAA	CAACCATGACTGTGTACCAAGAA	HTS of genomic target site
pSM067a51	ATGCGTGAGTGTGGATATGTG	CCTACATCACAGGAGGAAGGG	HTS of genomic target site
pSM067a52	CTGTCCTCCCTCAAGATACAGG	CCCCTTCCCTATGGAAATAATA	HTS of genomic target site
pSM067a53	TTCTTCCCTTCTGCTTCTGA	GCTTGGGACAACCATAACATCT	HTS of genomic target site
pSM067a54	CTCTTGCTCCACTGGTTGT	ATGTTCCAATCAGTACGCAGAG	HTS of genomic target site

pSM067a55	CGCAAGACAGTTGCAGAAGTA	TCTTGAGTAGTGAACAAGGGACTCT	HTS of genomic target site
pSM067a56r	AGTACTCTTCCAGACCCACGA	CTGGAAAGTCTCAGAACGCCCTA	HTS of genomic target site
pSM067a57	TGGGTCTTCATCAGAGAGTCCT	ACAACCAGTGGAGGCAAGAG	HTS of genomic target site
pSM067a58r	CTAGCCATCCTCTTGTCTCTGC	GGCCATTACAGAAGAAAGGAAA	HTS of genomic target site
pSM067a59	CTACTACCCCCATCTCTCCCTC	CATT CAGCTTCCCTTGCTTCTC	HTS of genomic target site
pSM067a60r	ATCCTCAAGAAGGAAGTCATCG	CTAGAGACAGCCCCAGGAGTC	HTS of genomic target site
pSM067a61	CTGCCTCCTGCTAGGTCTTAG	CCTTCGAGAAGAACCACTACG	HTS of genomic target site
pSM067a62	CATTCGGAGATTTGGATGTTCT	GCATTTCCACAGCTACACCATA	HTS of genomic target site
pSM067a63	AGAGAGAGCAGGACGTCACAGT	CAAGGTGAAAGCGGAAGTAGG	HTS of genomic target site
pSM067a64r	GCTATTATGAAGCCATTACCGC	CTGTTGGCTCTGGATTCTTC	HTS of genomic target site
pSM067b2	GCTTTAACACATTGTCTGTGCG	GT TTTCTGTC CCTCCCTCAGTA	HTS of genomic target site
pSM067b4	GTCTGGTGCCATGGAGAGTAG	GGTATCAGGCGACGTGGTAT	HTS of genomic target site
pSM067b5	CACTACATGCCAGCTACTTGC	AGCTCTCGTAGTGGTGCATTT	HTS of genomic target site
pSM067b7	CAGTCCCAGATGAGTGACAT	CTAAAGTATTTGACCTCGGGC	HTS of genomic target site
pSM067b9	CAGAGAGAGCAGGACGTCACA	AGCACTACCTACGTCAGCACCT	HTS of genomic target site
pSM067b12	CTACTACCCCCATCTCTCCCTC	CATT CAGCTTCCCTTGCTTCTC	HTS of genomic target site
pSM067b14	CTGACTCATGAGGGGACTTTA	CAAGACCTCATCAGGTGAGCAT	HTS of genomic target site
pSM067b15	CTTCAGGAAGTTGTTCTGAGGC	TGAGGAGACACTCCAAGAGAGC	HTS of genomic target site
pSM067b17	ACCACGTAGTGGTTCTCTCG	CTCTGTCTGTACCTCGGTGTGT	HTS of genomic target site
pSM067b19	CACTGCTGAACCAGTCAAACTC	GGCATGGGAAATATAAACTTG	HTS of genomic target site
pSM067b22	ATGCGTGAGTGTGGATATGTG	CCTACATCACAGGAGGAAGGG	HTS of genomic target site
pSM067b24	CACTACATGCCAGCTACTTGC	AGCTCTCGTAGTGGTGCATTT	HTS of genomic target site
pSM067b26	CCAGAGAGAAAGGAGAGGGAG	GACCCGATGCGGTTAGAG	HTS of genomic target site
pSM067b27	AGAAGCCAGTGGACTAGCACTT	GGACTCTCTGATGAAGACCCAG	HTS of genomic target site

pSM067b29	CACTACATGCCAGCTACTTG	AGCTCTCGTAGTGGTGCATT	HTS of genomic target site
pSM067b32	AGGGAAAGCCACTCACGAAG	GCGGCTTACCATAGAGTCCT	HTS of genomic target site
pSM067b33	GTAGAAATGGGTCTTGCTTG	TTGAGTCTATCGAGTGTGTCAT	HTS of genomic target site
pSM067b34	GACAGAGGGAGAGAACAGAGC	TTCTAGATGCCGACAAAAGGAT	HTS of genomic target site
pSM067b38	GTAGAAATGGGTCTTGCTTG	TTGAGTCTATCGAGTGTGTCAT	HTS of genomic target site
pSM067b39	ATCTCTTCAGCCCCTGAGTTGT	TTCGTGACCCTGAGTGTATGTG	HTS of genomic target site
pSM067b41	AGGGGTAGAAATGGAGAGGGTC	CTCAACTACCCTCACCCAG	HTS of genomic target site
pSM067b43	CTACTACCCCCATCTCTCCCTC	CATT CAGCTTCCCTTGCTTCTC	HTS of genomic target site
pSM067b45	TAAATT CCTGAAGCTCTCCAA	TGGAATGAATGGCTGAATTATG	HTS of genomic target site
pSM067b48	AGATGAAGAGACAAGCTGGGAG	ACTTTCAAATGGCTTCACCC	HTS of genomic target site
pSM067b49	CTACTACCCCCATCTCTCCCTC	CATT CAGCTTCCCTTGCTTCTC	HTS of genomic target site
HBB	ACACTCTTCCCTACACGACGCTCTC CGATCTNNNNAGGGTGGCCAATCTA CTCCC	TGGAGTT CAGACGTGTGCTTCCGATCTGTCTTCTCT GTCTCCACATGCC	HTS of genomic target site

Supplementary Table 9. Protospacer sequences used in this work.

name	purpose	gene (if applicable)	protospacer	PAM
G7'	PACE AP protospacer		AGTCTCCTCAGCAAAACGAA	multiple
VEGF2	PACE AP protospacer	VEGFA	GACCCCTCCACCCCGCTC	multiple
HEK3	nuclease selection and PAM depletion		GGCCCAGACTGAGCACGTGA	multiple
HEK4	GUIDE-seq		GGCACTGCGCTGGAGGTGG	GGG
pSM067a1	human genomic site target	AKT1	GGGACAGAGGAGCAAGGTTT	AAAT
pSM067a2r	human genomic site	TTN	GACTCTGCCAGCATCATGGT	AAAA
pSM067a3	human genomic site	EMX1	GGGGGACACTGGGATCACT	AAAC
pSM067a4	human genomic site	PIK3CA	GAGTGCACTATTATAACCC	AAAG
pSM067a5	human genomic site	IDH2	GGTCTTATAGTGCCTGGAG	TAAA
pSM067a6r	human genomic site	TTN	GTCTGAAGCATGAGTCGGT	TAAT
pSM067a7	human genomic site	PIK3CA	GTCCTGTACTCTGGATCTT	TAAC
pSM067a8	human genomic site	FANCF	GCACCACTACGAAGAGCTAA	TAAG
pSM067a9r	human genomic site	AKT1	GCCCAGCAGCTTCAGGTACT	CAAA
pSM067a10	human genomic site	AKT1	GCAGGACCTGCCGGCCCCC	CAAT
pSM067a11	human genomic site	IDH2	GTCAAGGAGTGGAAAGTGT	CAAC
pSM067a12	human genomic site	FANCF	GGCTGCACAACCAGTGGAGG	CAAG
pSM067a13	human genomic site	IDH2	GCCTCACGTACCATGAGGG	GAAA
pSM067a14r	human genomic site	TTN	GTCTATAACCTGTGAAGCCAA	GAAT
pSM067a15	human genomic site	AKT1	GTGGCCCACACACTCACCAG	GAAC
pSM067a16r	human genomic site	TTN	GAACCAGTAATTCAAGCAGTC	GAAG
pSM067a17	human genomic site	PIK3CA	GCACATCATGGTGGCTGGAC	AACA
pSM067a18	human genomic site	FANCF	TGGAAGTCGCTAATCCGG	AACT
pSM067a19	human genomic site	FANCF	GCTTCAATGGCTATAGAGAG	AACC
pSM067a20	human genomic site	EMX1	GACAGGCCAGGTCTTCTT	AACG
pSM067a21r	human genomic site	TTN	GTACTCTAACTGGCTCTGGG	TACA
pSM067a22	human genomic site	FANCF	GTCAATGCTTAAAGGGACA	TACT
pSM067a23	human genomic site	IDH2	GGAGCAGCGTCCTCCAGCG	TACC
pSM067a24	human genomic site	EMX1	GCAGCCGCCGCCTGGCCG	TACG
pSM067a25	human genomic site	PIK3CA	GACCGATTGCATAGGAATTG	CACA
pSM067a26	human genomic site	IDH2	GCCTACCACCCAGGCCACG	CACT
pSM067a27	human genomic site	AKT1	GCCTCAGTTCAACCTGGTGG	CACC
pSM067a28	human genomic site	FANCF	GCCCAC TGCAAGGCCGGCG	CACG
pSM067a29	human genomic site	AKT1	GGCCAGGGTTACCCAGTGG	GACA
pSM067a30r	human genomic site	TTN	GAGACAAGGAAGCCAAGTGA	GACT
pSM067a31r	human genomic site	TTN	GTTCAACGACTGCAGAGCAT	GACC
pSM067a32	human genomic site	IDH2	GGCCCAGGGTACGCTGGGAG	GACG
pSM067a33	human genomic site	IDH2	GGAAGCCTATGGCGTTGCAA	AATA
pSM067a34r	human genomic site	TTN	GCCCAC TTGATGCCAGATC	AATT
pSM067a35	human genomic site	PIK3CA	GACCAATTGGCATGCTCTTC	AATC
pSM067a36	human genomic site	PIK3CA	GCAGCACGAGGAAGATCAGG	AATG
pSM067a37	human genomic site	PIK3CA	GTCATGGTTGATTTCAGAG	TATA
pSM067a38	human genomic site	AKT1	GGCACCATGAGCGACGTGGC	TATT
pSM067a39	human genomic site	FANCF	GTACTTCCAATCACTTCCTC	TATC

pSM067a40	human genomic site	IDH2	GTTTACCTCAGCCAGTCAG	TATG
pSM067a41	human genomic site	AKT1	GGCTGACACAATCTCAGCGC	CATA
pSM067a42	human genomic site	PIK3CA	GATCTTAGTCACCCATGTAG	CATT
pSM067a43	human genomic site	IDH2	GGCCAGGATGTCTGACTGCA	CATC
pSM067a44	human genomic site	FANCF	GTAATCTACAAGATGGCCTT	CATG
pSM067a45	human genomic site	FANCF	GTAAACAAGGCCAAACTCCA	GATA
pSM067a46	human genomic site	FANCF	GCAGCTGTATCTCACACTAT	GATT
pSM067a47	human genomic site	AKT1	GTCAAGTGCTACCGTGGAGA	GATC
pSM067a48	human genomic site	PIK3CA	GTAACATCATGGTGAAAGAC	GATG
pSM067a49	human genomic site	IDH2	GGCCACACAAAGCCACCCG	AAGA
pSM067a50	human genomic site	PIK3CA	GTCAGTATAAGGTGAGTAAC	AAGT
pSM067a51	human genomic site	AKT1	GCCACACGATAACCGGCAAAG	AAGC
pSM067a52	human genomic site	EMX1	GGGTACTGGTGGAGGGGGTC	AAGG
pSM067a53	human genomic site	PIK3CA	GATCCCCCAAGAACATCTAG	TAGA
pSM067a54	human genomic site	FANCF	GTCCCAGGTGCTGACGTAGG	TAGT
pSM067a55	human genomic site	PIK3CA	GCCCTTAAGGGATCTAAA	TAGC
pSM067a56r	human genomic site	akt1	GTACGCCAACGGGGCGAGG	TAGG
pSM067a57	human genomic site	FANCF	GAAGCGCAGCATGTGCACCG	CAGA
pSM067a58r	human genomic site	IDH2	GTCTTCACCCAAAAGATGG	CAGT
pSM067a59	human genomic site	AKT1	GGGTACTAACCTCGTTGTG	CAGC
pSM067a60r	human genomic site	AKT1	GGCCACCTCGTCCTGTAAAG	CAGG
pSM067a61	human genomic site	EMX1	GGGTGATTACCTCGTCTCG	GAGA
pSM067a62	human genomic site	PIK3CA	GCACAAAGAACCTTATTCT	GAGT
pSM067a63	human genomic site	FANCF	GCCCCATTGCGACGGCTCTG	GAGC
pSM067a64r	human genomic site	TTN	GGACCAAAGAGTTAACTGAA	GAGG
pSM067b2	human genomic site	FANCF	GGTTAACAGGCCAAACTCC	AGAT
pSM067b4	human genomic site	AKT1	GGCGTACTCCATGACAAAGC	AGAG
pSM067b5	human genomic site	FANCF	GGCAATGCTGTGTTATTACT	TGAA
pSM067b7	human genomic site	IDH2	GGCCAGTGCAGAGTCATGG	TGAC
pSM067b9	human genomic site	FANCF	GGGTCCAGTTCCGGGATTAG	CGAA
pSM067b12	human genomic site	AKT1	GCCACGTCGCTCATGGTGC	CGAG
pSM067b14	human genomic site	IDH2	GCATCTGGCAAACCTATGG	GGAT
pSM067b15	human genomic site	FANCF	GAGAACCCAAATCTCAGGA	GGAC
pSM067b17	human genomic site	EMX1	GTGCGAAGGGGGCGTGCAGA	AGCA
pSM067b19	human genomic site	PIK3CA	GGTACTGGCAAAGATTCAA	AGCC
pSM067b22	human genomic site	AKT1	GCACCTTCTCTCGTACACG	TGCT
pSM067b24	human genomic site	FANCF	GCCTATACAGAACTGAGGCC	TGCG
pSM067b26	human genomic site	PIK3CA	GCCAGTCCGCGCTACTCA	CGCT
pSM067b27	human genomic site	FANCF	GACAAAGGCGGCTGCAACAG	CGCC
pSM067b29	human genomic site	FANCF	GCAGGCCTCAGTTCTGTATA	GGCA
pSM067b32	human genomic site	EMX1	GAGGAACCGCTCCGGCCAC	GGCG
pSM067b33	human genomic site	PIK3CA	GACAATGTGAAACACTCAAAG	AGTA
pSM067b34	human genomic site	IDH2	GATGACCGTATTATCTGGC	AGTT
pSM067b38	human genomic site	PIK3CA	GGAACAAGGTACTCTTGAG	TGTT
pSM067b39	human genomic site	AKT1	GCCACAGTCTGGATGGCGGT	TGTC
pSM067b41	human genomic site	EMX1	GAACACGAGCTCGGGCCCAC	CGTA

pSM067b43	human genomic site	AKT1	GCCTGAGAGGAGCGCGTGAG	CGTC
pSM067b45	human genomic site	PIK3CA	GGCCGAAAGGGTGCTAAAGA	GGTA
pSM067b48	human genomic site	IDH2	GGCATGTACAACACCGACGA	GGTG
pSM067b49	human genomic site	AKT1	GAGCGACGTGGCTATTGTGA	AGGA
HbS-install	install HbS mutation	HBB	GTAACGGCAGACTTCTCCTC	AGG
HbS-CATG	human genomic site	HBB	TCCTCAGGAGTCAGGTGCAC	CATG
HbS-CACC	human genomic site	HBB	TTCTCCTCAGGAGTCAGGTG	CACC

Supplementary Note 1. Evolution of SpCas9 variants that recognize NAT or NAC PAM sequences.

Based on the outcomes of the NAA PAM evolution campaigns, we approached the evolution of SpCas9 variants capable of recognizing NAT and NAC PAM sites in a manner that minimized potentially deleterious mutations. To ensure that we started with nuclease-active variants, we developed a modified version of a previously reported bacterial DNA cleavage selection (**Supplementary Figures 1e,f**)³. In this nuclease selection, SpCas9 variants are challenged for their ability to cleave a protospacer-PAM sequence on a high-copy plasmid that also encodes a conditionally toxic gene (*sacB*). The surviving cells encode nuclease-active SpCas9 variants that cleave the target sequence, destroying the toxic plasmid.

Thus, we converted the dSpCas9 clones from the NAT or NAC PANCE pools into nuclease-active forms by restoring Asp 10 and His 840, then passed the resulting libraries through the nuclease selection using a TAT or CAC PAM, respectively. We isolated two clones (SacB.TAT-1 and -2; **Fig. 2e**) that exhibited DNA cleavage activity on the TAT PAM with PID consensus mutations of D1135N, E1219V, Q1221H, P1321S, and R1335L, and a third clone that cleaved a CAC PAM with PID mutations N1135D, E1219V, D1332N, R1335Q, and T1337N (SacB.CAC; **Fig. 2f** and **Supplementary Table 2**). We evolved these nuclease-active TAT and CAC variants using split-dSpCas9 PACE in host cells encoding APs with either NAT (AAT or TAT) PAMs or NAC (AAC, TAC, or CAC) PAMs, respectively. These experiments resulted in the enrichment of several additional PID mutations, including R1114G, which arose independently in all three trajectories (NAA, NAT, and NAC) (**Fig. 2b, e, and f** and **Supplementary Table 2**), suggesting that this mutation may be generally beneficial for modifying PAM recognition by the PID in a manner compatible with NA PAMs.

Next, we removed gVI from these evolved SP pools and subjected them to additional selection in split-dSpCas9 PACE using the dual-AP system (**Fig. 2a** and **Supplementary Fig. 1c**). Both protospacers contained either an AAT or TAC PAM for evolution following the NAT or NAC trajectory, respectively. Increasing stringency for the NAT-targeting SpCas9 did not improve activity despite enrichment of several mutations (TAT.P6; **Fig. 2d** and **Supplementary Fig. 2a**). We therefore selected the most active NAT PAM-targeting variant from the split-dSpCas9 evolution (TAT.P5-1; Figure 2D) to move forward with. This variant contained the 11 PID mutations R1114G, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, E1253K, P1321S, D1332G, R1335L (**Fig. 2e, g** and **Supplementary Table 2**). PACE of NAC-targeting split-dSpCas9 using dual protospacers and a TAC PAM also enriched for

several mutations (TAC.P9; **Fig. 2g**). We shuffled residues 574-1368 of the surviving clones with that of SpCas9 and re-challenged the resulting library with our most stringent binding selection (TAC.P9s; **Fig. 2g**). From the surviving SP pool, we isolated clone TAC.P9s-3 with the PID mutations R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N, S1338T, and H1249R (**Fig. 2f, h**, and **Supplementary Table 2**).

Supplementary Note 2. Bacterial PAM depletion

As an initial characterization of the PAM compatibility profile of these three evolved variants as nucleases, we used a bacterial PAM depletion assay³ on an NNNN PAM library. For comparison, we also performed PAM depletion experiments with SpCas9-NG in parallel. Cells were plated after 1-hour, 3-hour, or overnight expression of the SpCas9 variant to assess kinetic differences in PAM preference (**Supplementary Fig. 3**). Depletion scores of any given PAM (defined as the frequency of the PAM in the input library divided by the frequency of the PAM post-selection) increased with longer induction times as expected (**Supplementary Table 3**), with overnight inductions showing more promiscuous PAM compatibility for both G and A at the second PAM base for all variants tested, including SpCas9-NG (**Supplementary Fig. 3**). In contrast, the shortest induction times yielded the clearest PAM preferences, with SpCas9-NRRH, -NRTH, and NRCH exhibiting a third PAM base preference consistent with the PAM each was evolved to recognize (A, T, or C, respectively), and a fourth PAM base preference for H (**Supplementary Fig. 3**). Interestingly, SpCas9-NG also displayed a moderate preference for G at the third and fourth PAM bases at short induction times (**Supplementary Fig. 3**).

Supplementary Note 3. sgRNA preferences of evolved variants and SpCas9-NG in mammalian cells.

Our human cell target site library also allowed us to investigate the tolerance of our variants to changes in protospacer length and the presence of 5' mismatches between the sgRNA protospacer and the target DNA sequence. The commonly used human U6 promoter initiates transcription with a 5' G. Guide sequences are typically either extended to the next native G or transcribed with a mismatched 5' G at position -1 if a G is not natively present at the 5' end of the protospacer. However, high-fidelity (HF) SpCas9s⁴⁻⁸ generally exhibit decreased efficiency when using a 21 nucleotide (nt) guide with a mismatched 5' G⁹. Because PACE has previously led to SpCas9s with HF properties⁸, we sought to determine if our new

variants shared the same characteristics by investigating the average base editing activity of our evolved variants across all 11,776 library sites containing either a 20 nt protospacer with a matched 5' G ("20-matched"), a 21 nt protospacer with a matched 5' G ("21-matched"), or a 21 nt protospacer with a mismatched 5'G ("21-mismatched").

The three SpCas9 variants evolved in this study and SpCas9 all showed the highest base editing activity with a 20-matched sgRNA (Supplementary Figure 4C-D; interestingly, however, SpCas9-NG performed best with a 21-matched sgRNA (**Supplementary Fig. 4b**). When examining all NRNN PAMs, our variants and SpCas9 also showed a significant decrease in base editing efficiency when the sgRNA protospacer was increased to 21 nt, regardless if the 5' G was matched with the target sequence (**Supplementary Fig. 4b-c**); in contrast, for SpCas9-NG this was only true when the 21-mismatched sgRNA (**Supplementary Fig. 4c**). The magnitude of this decrease was similar to or greater for our evolved variants (SpCas9-NRRH: $23\pm2.7\%$, SpCas9-NRTH: $12\pm2.9\%$, SpCas9-NRCH: $14\pm2.9\%$) when compared to SpCas9 ($13\pm5.3\%$). In contrast, SpCas9-NG demonstrated a preference for 21-matched sgRNAs, leading to an average $18.5\pm5.4\%$ increase of editing efficiency when compared to 20-matched sgRNAs (**Supplementary Fig. 4b**); however, a decrease in editing efficiency was still observed with 21-mismatched sgRNAs ($7.3\pm3.2\%$, **Supplementary Fig. 4c**). Interestingly, the deleterious effect of using a 21 nt protospacer on the editing efficiency of our evolved variants and SpCas9 is lessened when targeting sites with NGNN or NGGN PAMs (**Supplementary Fig. 4b-c**). Together, these results suggest that our evolved variants are somewhat sensitive to the use of 21-nt sgRNA protospacers, and that this sensitivity is exacerbated by the presence of 5' G mismatches. Additionally, these experiments suggest that the optimal sgRNA protospacer length for SpCas9-NG may be longer than 20 nt.

Supplementary Note 4. Evolved SpCas9 amino acid sequences.

SpCas9-NRRH:

MDKKYSIGLTIGTNSVGWAVITDEYKPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAETRLK
RTARRRYTRRKNRICYLQEIFSNEAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLVDSTDKADLRILYLALAHMIKFRGHFLIEGDLNPNSDVKLFQLVQTYNQ
LFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSGLTPNFKSNFDLA
EDAQLQLSKDTYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMVKR
YDEHHQDLTLLKALVRQQLPEKYKEIFFDQSCKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEE
LLVQLNREDLLRKQRTFDNGIIPHQIHLGELHAILRRQGDFYPFLKDNRKVNTEITKAPLSASMVKR
LARGNSRFAMTRKSEETITPWNFEEVVDKGASAQSFIERNFNDKLPNEKVLPKHSLLYE
YFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI
SGVEDRFNASLGTYHDLLKIJKDFLDNEENEDILEDIVLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRLRYTGWRGLSRKLINGIRDQKSGKTILDFLSDGFANRNFMQLIHDDSLTFKEDIQK
AQVSGQGDSLHEHIANLAGSPAICKGILQTVKVVDELVKVMGGHKPENIVIEMARENQTTQKG
QKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYLQNNGRDMDYVDQELDINRLSDY
DVDHIVPQSFNLDKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDLSRMNTKYDENDKLIREVKVITLKSCLV
SDFRKDFQFYKVRREINNYHHAHDAYLNAVVTALIKKYPKLESEFVYGDYKVDVRKMIAKSE
QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMP
QVNIVKKTEVQTGGFSKESILPKGNSDKLIARKKDWPKKYGGFNSPTAAYSVLVAKVEKGK
SKKLKSVKELLGITIMERSSFEKPNIGFLEAKGYKEVKKDLIILPKYSLFELENGRKRLMASAG
VLHKGNEALPSKYVNFLYLAHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILA
DANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGVPAAFKYFDTTIDKKRYTSTKEVLDATLH
QSITGLYETRIDLSQLGGD

SpCas9-NRTH:

MDKKYSIGLTIGTNSVGWAVITDEYKPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAETRLK
RTARRRYTRRKNRICYLQEIFSNEAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLVDSTDKADLRILYLALAHMIKFRGHFLIEGDLNPNSDVKLFQLVQTYNQ
LFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSGLTPNFKSNFDLA
EDAQLQLSKDTYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMVKR
YDEHHQDLTLLKALVRQQLPEKYKEIFFDQSCKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEE
LLVQLNREDLLRKQRTFDNGIIPHQIHLGELHAILRRQGDFYPFLKDNRKVNTEITKAPLSASMVKR
LARGNSRFAMTRKSEETITPWNFEEVVDKGASAQSFIERNFNDKLPNEKVLPKHSLLYE
YFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI
SGVEDRFNASLGTYHDLLKIJKDFLDNEENEDILEDIVLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRLRYTGWRGLSRKLINGIRDQKSGKTILDFLSDGFANRNFMQLIHDDSLTFKEDIQK
AQVSGQGDSLHEHIANLAGSPAICKGILQTVKVVDELVKVMGGHKPENIVIEMARENQTTQKG
QKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYLQNNGRDMDYVDQELDINRLSDY
DVDHIVPQSFNLDKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDLSRMNTKYDENDKLIREVKVITLKSCLV
SDFRKDFQFYKVRREINNYHHAHDAYLNAVVTALIKKYPKLESEFVYGDYKVDVRKMIAKSE
QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMP
QVNIVKKTEVQTGGFSKESILPKGNSDKLIARKKDWPKKYGGFNSPTVAYSVLVAKVEKGK
SKKLKSVKELLGITIMERSSFEKPNIGFLEAKGYKEVKKDLIILPKYSLFELENGRKRLMASASV
LHKGNEALPSKYVNFLYLAHYEKLKGSSEDNKQKQLFVEQHKHYLDEIIEQISEFSKRVILAD
ANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGVASAFAFKYFDTTIGRKLYTSTKEVLDATLH
SITGLYETRIDLSQLGGD

SpCas9-NRCH:

MDKKYSIGLTIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAATRLK
RTARRRYTRRKNRICYLQEIFSNEAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLVSTDKAIDLRLIYLALAHMIKFRGHFLIEGDLNPNSDVKLFQLVQTYNQ
LFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSGLTPNFKSNFDLA
EDAQLQLSKDTYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMVKR
YDEHHHQDLTLLKALVRQQQLPEKYKEIFFDQSNGYAGYIDGGASQEEFYKFIKPILEKMDGTEE
LLVKLNRDLLRKQRTFDNGIIPHQIHLGELHAILRRQGDYPFLKDNRKIEKILTFRIPYYVGP
LARGNSRFAMTRKSEETITPWNFEEVVDKGASAQSFIERNMTNFDKKNLPNEKVLPKHSLLYE
YFTVYNELTKVVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI
SGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLFEDREMIEERLKTYAHLFDDK
VMKQLKRLRYTGWRGLSRKLINGIRDQKSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQK
AQVSGQGDSLHEHIANLAGSPAICKGILQTVKVVDELVKVMGGHKPENIVIEMARENQTTQKG
QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDY
DVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLV
SDFRKDFQFYKVREINNYHHAHDAYLNAVVTALIKKYPKLESEFVYGDYKVDVRKMIAKSE
QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMP
QVNIVKKTEVQTGGFSKESILPKGNSDKLIARKKDWDPKYGGFNSPTVAYSVLVAKVEKGK
SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRM LASAG
VLQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILA
DANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTINRKQYNTTKEVLDATLI
RQSITGLYETRIDSQLGGD

Supplementary References

1. Ringquist, S. *et al.* Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. *Mol. Microbiol.* **6**, 1219–1229 (1992).
2. Davis, J. H., Rubin, A. J. & Sauer, R. T. Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res.* **39**, 1131–1141 (2011).
3. Kleinstiver, B. P. *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).
4. Chen, J. S. *et al.* Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* **550**, 407–410 (2017).
5. Kleinstiver, B. P. *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
6. Slaymaker, I. M. *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science (80-.).* **351**, 84–88 (2016).
7. Lee, J. K. *et al.* Directed evolution of CRISPR-Cas9 to increase its specificity. *Nat. Commun.* **9**, 3048 (2018).
8. Hu, J. H. *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018).
9. Zhang, D. *et al.* Perfectly matched 20-nucleotide guide RNA sequences enable robust genome editing using high-fidelity SpCas9 nucleases. *Genome Biol.* **18**, 191 (2017).